

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: September 26, 2006, 00:11:50 / Search time 7.9 Seconds  
(without alignments)  
2871.888 Million cell updates/sec

Title: US-10-537-002-7

Perfect score: 1560  
Sequence: 1 atgagcgcgtactgctgtca.....ccaagcagctatgtgta 786

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 283416 segs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seg length: 0  
Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 200 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=xlh  
-Q=/abss/ABSSWEB.spool/US10537002/runat\_25092006.170505.8414/app.query.fasta\_1  
-DB=FIR -QFMT=fastan -SUPERX=n2p.rpr -MINMATCH=0.1 -LOOCL=0 -LOOEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human4.0.cdi -LIST=200  
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL  
-OUTMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss802h  
-USER=US10537002 -CGEN\_1\_1\_25 -@runat\_25092006.170505.8414 -MCP=6 -ICPU=3  
-NO MWAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

FIR\_80: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	297	19.0	280	2	A39484 androgen-wthdrawa
2	135	8.7	173	2	A49182 integrin membrane
3	133	8.5	173	2	A48300 lens membrane prot
4	131.5	8.4	157	2	G02355 tumor-associated m
5	126.5	8.1	160	2	S21721 tumor-arrest-spec
6	121.5	7.8	160	2	JC5730 tumor-associated m
7	120.5	7.7	160	2	JN0503 peripheral myelin
8	118.5	7.6	1794	2	T38459 hypothetical diver
9	117.5	7.5	160	2	A41144 growth arrest-rela
10	114	7.3	695	2	G64327 H+-transporting tw
11	107	6.9	167	2	JC5044 epithelial membran
12	104.5	6.7	735	2	J45059 hypothetical prote
13	103.5	6.6	593	1	GYHU granulatin precursor
14	102.5	6.6	448	2	G53610 ntup protein - Ent

15	102.5	6.6	466	2	AD3550
16	102.5	6.6	1574	2	T13954
17	101.5	6.5	530	2	A45690
18	101	6.5	331	2	C95998
19	101	6.5	873	2	A47283
20	101	6.5	1964	2	T09059
21	100.5	6.4	2318	2	S45306
22	100.5	6.4	2318	2	S45306
23	100	6.4	411	2	PC2061
24	100	6.4	492	2	S41288
25	100	6.4	1007	2	T01437
26	99	6.3	521	2	S74569
27	99	6.3	865	2	A47282
28	98.5	6.3	722	2	S57246
29	98	6.3	712	2	AD2721
30	98	6.3	714	2	H97502
31	98	6.3	2187	2	T30826
32	97.5	6.2	1964	2	T09059
33	97.5	6.2	2825	2	T14271
34	96.5	6.2	144	2	A36324
35	96.5	6.2	292	2	A33524
36	96.5	6.2	775	1	EDBE11
37	96	6.2	167	2	JC5732
38	96	6.2	230	2	A38346
39	96	6.2	331	2	A33140
40	96	6.2	331	2	E98147
41	96	6.2	361	2	T45918
42	96	6.2	411	2	PC2060
43	96	6.2	472	2	B87658
44	96	6.2	595	2	B97415
45	96	6.2	595	2	AF2632
46	96	6.2	640	2	J01584
47	96	6.2	927	2	T00357
48	95.5	6.1	446	2	S01187
49	95.5	6.1	446	2	I25797
50	95.5	6.1	691	2	A25704
51	95.5	6.1	2531	2	A51888
52	95	6.1	329	2	S01075
53	95	6.1	341	2	A13425
54	95	6.1	383	2	D72655
55	95	6.1	476	2	B97661
56	95	6.1	481	2	BD4463
57	95	6.1	2321	1	KGZOHF
58	94.5	6.1	221	1	A82766
59	94.5	6.1	264	2	A82766
60	94.5	6.1	463	2	T16810
61	94.5	6.1	4135	2	T42629
62	94	6.0	884	2	A31928
63	93	6.0	223	2	P09554
64	93	6.0	235	2	G85803
65	93	6.0	337	2	G95915
66	93	6.0	895	2	S20582
67	93	6.0	2555	2	A40043
68	92.5	5.9	572	2	T29880
69	92.5	5.9	1520	2	T00273
70	92.5	5.9	3020	2	A43932
71	92	5.9	313	2	JC5342
72	92	5.9	320	2	A95961
73	92	5.9	387	2	B49175
74	92	5.9	539	2	D69748
75	92	5.9	718	2	T02310
76	92	5.9	749	2	T38488
77	91.5	5.9	314	2	C84116
78	91.5	5.9	438	2	T11889
79	91.5	5.9	524	2	S88539
80	91.5	5.9	548	2	A70780
81	91.5	5.9	667	2	F70682
82	91.5	5.9	879	2	E91031
83	91.5	5.9	1116	2	A48292
84	91.5	5.9	1510	2	T13634
85	91.5	5.9	3635	2	T10053
86	91	5.8	349	2	A10274
87	91	5.8	388	2	AH1183

NAD(P) transhydrog  
MEGF6 protein - ra  
transactivator EN  
probable sugar upt  
calphostin - fruit  
notch4 - mouse  
notch 3 protein -  
notch 3 protein -  
genome polypept  
genome polypept  
hypothetical prote  
NADH2 dehydrogenas  
calcium-binding pr  
ventral nervous sy  
H+ translocating p  
h+ translocating p  
nascent polypeptid  
notch4 - mouse  
Doc4 protein, stre  
growth arrest-spec  
high-affinity bran  
immediate-early pr  
tumor-associated m  
ultra-high-sulfur  
hypothetical prote  
hypothetical prote  
genome polypept  
NAD(P) transhydrog  
hypothetical prote  
conserved hypotet  
genome polypept  
NADH2 dehydrogenas  
NADH2 dehydrogenas  
synapsin I - rat  
ribosome transp  
L-arabinose transp  
ribosome transp  
hypothetical prote  
probable membrane  
conserved hypotet  
notch3 protein - h  
histidine/alanine-  
membrane protein X  
probable integral  
tenascin-X - bovin  
glucose transport  
hypothetical prote  
hypothetical prote  
probable sugar ABC  
dyctrophin-associa  
notch protein homo  
hypothetical prote  
hypothetical prote  
mucin 2 precursor,  
Na+/H+ antiporter  
probable sugar upt  
Moch A protein -  
amino acid transpo  
hypothetical prote  
troponin - human  
ribose ABC transpo  
hypothetical prote  
disintegrin-like m  
probable membrane  
probable membrane  
probable outer mem  
mucin, tracheobron  
probable minor tai  
lamunin alpha 5 ch  
L-arabinose transp  
antibiotic resista

88	91	5.8	399	2	AB0155	probable drug resi
89	91	5.8	400	2	C97200	uncharacterized co
90	91	5.8	431	2	AF0381	probable ammonium
91	91	5.8	513	2	AA4150	structural protein
92	91	5.8	861	2	AA4825	Notch homolog Mofc
93	91	5.8	1700	2	S08167	Balbiant ring 3 pr
94	91	5.8	4006	2	T09070	probable tenascin
95	90.5	5.8	425	2	AG0055	hypothetical prote
96	90.5	5.8	431	2	B87285	phosphate ABC tran
97	90.5	5.8	432	2	B87190	potassium uptake p
98	90.5	5.8	454	2	F83760	hypothetical prote
99	90.5	5.8	455	2	AI1995	amino acid transpo
100	90.5	5.8	472	2	T34748	transmembrane tran
101	90.5	5.8	483	2	AG1226	potassium uptake p
102	90.5	5.8	483	2	G86072	probable efpa proc
103	90.5	5.8	530	2	A70589	MFS permease (limp
104	90.5	5.8	542	2	AF2587	hypothetical prote
105	90.5	5.8	542	2	F97369	mitochondrial nico
106	90.5	5.8	1068	2	T48756	hypothetical prote
107	90	5.8	182	2	T19126	cytochrome O ubiq
108	90	5.8	320	2	S72756	probable protochme
109	90	5.8	321	2	H86981	branched-chain am
110	90	5.8	356	2	G75444	probable response
111	90	5.8	448	2	G95416	probable MFS trans
112	90	5.8	456	2	G83449	aromatic amino aci
113	90	5.8	457	1	ORBCAA	NADH2 dehydrogen
114	90	5.8	736	2	T13225	histidine/alanine-
115	89.5	5.7	218	2	B25942	hypothetical prote
116	89.5	5.7	300	2	H84194	probable membrane
117	89.5	5.7	317	2	F96656	hypothetical prote
118	89.5	5.7	472	2	S28286	cytochrome-c oxida
119	89.5	5.7	488	1	H64537	protein C38C10.2 [
120	89.5	5.7	493	2	G88553	MG66 protein - ra
121	89.5	5.7	1574	2	T13954	notch-1 protein -
122	89.5	5.7	2531	2	AA6019	ribose ABC transpo
123	89	5.7	325	2	D98292	ABC transporter, m
124	89	5.7	325	2	AE2991	multidrug resistan
125	89	5.7	371	2	D69415	regulatory protein
126	89	5.7	372	2	T45524	hypothetical prote
127	89	5.7	445	2	T31898	histidine permease
128	89	5.7	475	2	E69643	probable sulfate p
129	89	5.7	495	2	F95850	hypothetical prote
130	89	5.7	603	2	T45093	genome polypotein
131	89	5.7	782	2	SI8031	hypothetical prote
132	89	5.7	795	2	T52516	hypothetical prote
133	89	5.7	1737	2	T00209	MG66 protein - hu
134	89	5.7	1786	1	MMMSB1	laminin beta-1 cha
135	89	5.7	3011	1	GNMVCH	genome polypotein
136	89	5.7	3566	1	AA0701	tenascin-X precurs
137	88.5	5.7	110	2	D72701	hypothetical prote
138	88.5	5.7	661	1	VCWVCB	env polypotein -
139	88.5	5.7	2471	2	AA9128	cell-fate determin
140	88	5.6	190	2	A71003	hypothetical prote
141	88	5.6	264	2	T10468	histidine/alanine-
142	88	5.6	333	2	AI0050	probable ABC trans
143	88	5.6	373	2	E49094	methylnalonyl-CoA
144	88	5.6	397	2	G90667	probable transcrip
145	88	5.6	397	2	C85518	hypothetical prote
146	88	5.6	411	2	T20623	phenylalanine-spec
147	88	5.6	458	2	AA8556	phenylalanine-spec
148	88	5.6	458	2	F90705	phenylalanine-spec
149	88	5.6	458	2	A39431	phenylalanine tran
150	88	5.6	492	2	G75389	NADH2 dehydrogen
151	88	5.6	502	2	B70845	probable eygar tra
152	88	5.6	543	2	D83262	hypothetical prote
153	88	5.6	553	2	F75407	probable transpor
154	88	5.6	579	2	T45344	hypothetical prote
155	88	5.6	664	2	T33959	probable urea acti
156	88	5.6	729	2	T13227	NADH2 dehydrogen
157	88	5.6	752	2	B82146	Rec2-related prote
158	88	5.6	784	2	J00317	hypothetical 82K p
159	88	5.6	1125	2	B41206	microtubule-associ
160	88	5.6	2531	2	T31070	notch homolog - se

161	87.5	5.6	303	2	AB3201	conserved hypothet
162	87.5	5.6	306	2	S08607	chorion protein s3
163	87.5	5.6	330	2	F64905	probable sugar tra
164	87.5	5.6	330	2	G85723	probable transport
165	87.5	5.6	330	2	B90894	probable transport
166	87.5	5.6	351	1	KGZQHL	histidine-rich gly
167	87.5	5.6	436	2	AA3145	invariant surface
168	87.5	5.6	443	2	H83589	conserved hypothet
169	87.5	5.6	449	2	AB2887	MFS permease (limp
170	87.5	5.6	472	2	F82639	resistance protein
171	87.5	5.6	477	2	G97662	benzoate transport
172	87.5	5.6	483	2	AE0914	lirk system potassi
173	87.5	5.6	509	2	G96766	protein lipase F2P
174	87.5	5.6	597	1	NBRUC4	Ca4-binding protei
175	87.5	5.6	719	2	T00266	hypothetical prote
176	87.5	5.6	741	2	T33791	NADH2 dehydrogen
177	87.5	5.6	770	2	T23999	hypothetical prote
178	87.5	5.6	784	2	V00317	hypothetical prote
179	87.5	5.6	1041	2	T15521	laminin beta-2 cha
180	87.5	5.6	1801	1	MMKRS	laminin alpha-1 ch
181	87.5	5.6	3075	2	SI4458	conserved hypothet
182	87	5.6	254	2	E69853	hypothetical prote
183	87	5.6	268	2	S34364	hypothetical prote
184	87	5.6	395	2	AB1061	probable inner me
185	87	5.6	413	2	B83180	probable MFS trans
186	87	5.6	420	2	AE2856	conserved hypothet
187	87	5.6	457	2	D85494	aromatic amino aci
188	87	5.6	457	2	D90643	aromatic amino aci
189	87	5.6	469	1	S29126	propterlin precuro
190	87	5.6	522	2	S41819	nucleoporin p62 -
191	87	5.6	600	2	T18593	hypothetical prote
192	87	5.6	2406	2	AA5148	od2 protein - frui
193	87	5.6	2515	1	S47008	tenascin-like prot
194	87	5.6	3011	2	SA0770	genome polypotein
195	86.5	5.5	166	2	C72641	hypothetical prote
196	86.5	5.5	346	2	AA3563	ribose transport s
197	86.5	5.5	397	2	G69295	oxalate/formate an
198	86.5	5.5	400	2	AE0808	nucleoside permeas
199	86.5	5.5	401	2	D83363	conserved hypothet
200	86.5	5.5	475	2	A70320	nitrate transpote

## ALIGNMENTS

## RESULT 1

androgen-withdrawal apoptosis protein RVp1, prostatic - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 03-Aug-1992 #sequence\_revision 03-Aug-1992 #text\_change 04-Mar-2000

C/Accession: A39484

R.Brüehl, M.M.; Miesfeld, R.L.

Mol. Endocrinol. 5, 1381-1388, 1991

A>Title: Isolation and characterization of transcripts induced by androgen withdrawal and

A/Reference number: A39484; MUID:92130987; PMID:1723140

C/Accession: A39484

A/Molecule type: mRNA

A/Residues: 1-280 <BRI>

A/Cross-references: UNIPARC:UPI0000163BB3; GB:W74067; NID:9205857; PIDD:AAA41760.1; PID:5

A/Genes: RVP.1

C/Superfamily: rat androgen-withdrawal apoptosis protein RVp1

## Alignment Scores:

Pred. No.:	2,786-16	Length:	280
Score:	297.00	Matches:	75
Percent Similarity:	47.8%	Conservative:	97
Best Local Similarity:	29.4%	Mismatches:	36
Query Match:	19.0%	Indels:	8
DB:	2	Gaps:	8

US-10-537-002-7 (1-786) x A39484 (1-280)

QY 28 GGATTCGGTTCACATGATGGATGGGACATCATTCGTGCACCTGATGACGACG 87

Db 9 GlyThrSerLeuAlaValLeuGlyTrpLeuGlySerThrIleValCysCysAlaLeuPromet 28  
 Qy 88 TGAAGACCCAGAGACTTGTACACAAACCC--GTACAGCTGTTTCACTACACAGGCG 144  
 Db 29 TrpArgValSerAlaPheIleGlySerSerIleIleThrAlaGlnIleThrTrpIleGly 48  
 Qy 145 CTGTGGCCCTCTGTGTCTGAGAGCTGTGCTTACCGAGTGTCCGGGGCTACTTCAAC 204  
 Db 49 LeuTrpMetAlaCysVal--GlnSerThrGlyGlnMetGlnCysValMetCysAspSer 67  
 Qy 205 CTGTGGGGCTGCTGAGGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 264  
 Db 68 LeuLeuAlaLeuProGlnAspLeuGlnAlaAlaGlnAlaLeuIleValValSerIleLeu 87  
 Qy 265 CTGGGTCCTATGGCTCTGCTGTATCATCTTGTCCGTAATGATCCGACATTGGACG 324  
 Db 88 LeuAlaAlaPheGlyLeuLeuValAlaLeuValGlyAlaGlnCysThrAsnCys--Val 106  
 Qy 325 ATGAGAGACTCTGCAAGCCCAACATGACATGACCTCCGAGATCATGTCTTCTTCA 384  
 Db 107 GlnAspGlnTrpAlaValAlaValIleThrIleValAlaGlyValLeuPheLeuAla 126  
 Qy 385 GGTCTTGTGCAATGCTGAGAGTGTGTGTGCTGCTGCAATGCTGCTGCTGCTGCTG 444  
 Db 127 AlaValLeuThrLeuValAlaProValSerTrpSerAlaSerThrIleIleArgAspPheTyr 146  
 Qy 445 ATGTCCACAGCTAACATGATGTACACCGGATGGTGGATGGTGGATGGTGGATGG 504  
 Db 147 TACACATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 156  
 Qy 505 TACACATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 564  
 Db 157 --GluMetGlyThrGlyLeuValGlyIleThrAlaAlaAlaLeuGlnLeuLeuGly 175  
 Qy 565 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 615  
 Db 176 GlyAlaLeuLeuGlyCysSerCysProProArgGlnGlyTyrAlaProThrIleValLeu 195  
 Qy 616 TACAAAGCCGTTTCTTATCATGCTCAGGCGCAAGTGTGCTCCTACAGCTGGAGGCTTC 675  
 Db 196 TyrSerAla----- 198  
 Qy 676 AAGGCCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 735  
 Db 199 ProArgSerThrGlyProGlyThrGlyThrGlyThr-----AlaTyrAspArgGlyThr 216  
 Qy 736 CGCACAGAGACAGAGGTAACAATCTTATCCTTCAGACAGACTAT 780  
 Db 217 ThrSerGlnArgProGlyAlaArgThrProHisHisHisTyr 231  
 RESULT 2  
 A49182  
 N:Alternate names: lens membrane protein MP20 - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
 C/Accession: A49182; S27883  
 R:Kumar, N.M.; Javali, L.J.; Tenbroek, E.; Louis, C.F.  
 Exp. Eye Res. 56, 35-43, 1993  
 A>Title: Cloning and expression of a major rat lens membrane protein, MP20.  
 A/Reference number: A49182; MUID:93162128; PMID:7679355  
 A/Accession: A49182  
 A:Status: Preliminary  
 A:Molecule type: nucleic acid  
 A:Residues: 1-173 <KUN>  
 A:Cross-references: UNIPROT:P54825; UNIPARC:UPI000012E77B; GB:S55224; NID:g265405; PIDN:  
 A:Experimental source: lens  
 A>Note: sequence extracted from NCBI backbone (NCBIN:124760, NCBI:P124761)  
 R:Kumar, N.M.; Javali, L.J.; Tenbroek, E.; Louis, C.F.  
 Submitted to the EMBL Data Library, February 1992  
 A>Description: Cloning and expression of a major lens membrane protein, MP20.  
 A/Reference number: S27883

A/Accession: S27883  
 A:Molecule type: mRNA  
 A:Residues: 1-173 <KUN2>  
 A:Cross-references: UNIPARC:UPI000012E77B; EMBL:M87053; NID:g205514; PIDN:AAA41631.1  
 C:Superfamily: growth arrest-specific protein  
 C:Keywords: membrane protein  
 Alignment Scores:  
 Pred. No.: 0.00459 Length: 173  
 Score: 135.00 Matches: 48  
 Percent Similarity: 38.6% Conserved: 30  
 Best Local Similarity: 23.8% Mismatches: 78  
 Query Match: 8.7% Indels: 46  
 DB: 2 Gaps: 9  
 US-10-537-002-7 (1-786) x A49182 (1-173)  
 Qy 22 GGTGGGGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 81  
 Db 6 GlyGlyGlyLeuPhePheCysAlaIleTrpValGlyIleThrIleLeuLeuValAlaThrAlaThr 25  
 Qy 82 GACCAAGAGACACCCAAAGACTTGTACAAACCCGTACAGCTGTTTCACTACAG 141  
 Db 26 AsnIleTrpMetGlnTyrArgLeu-----SerGlySerPheAlaHisGln 40  
 Qy 142 GGGCTGTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 198  
 Db 41 GlyLeuTrpArgTyrCysLeuGlyValValCysPheLeuGlnThrGlnSerIleAlaTyr 60  
 Qy 199 TTCACCTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 258  
 Db 61 Trp-----AsnAlaThrArgAlaPheMetIleLeuSer 71  
 Qy 259 ATGCTCTGGGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 318  
 Db 72 AlaLeuCysAlaThrSerGlyIleIleMetGlyAlaLeuAla-----Phe 86  
 Qy 319 GGCAGCAGAGAGACTGCTGCAAGCCCAACATGACATGACCTGCGGATCATGTTTCA 378  
 Db 87 AlaGlnIleSerThrPheThrArgLeuSerArgProPheSerAlaGlyIleMetPhePhe 106  
 Qy 379 GTCCTAGGCTTGTGCAATGCTGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 438  
 Db 107 AlaSerThrLeuPheValLeuAlaLeuAlaIle----- 118  
 Qy 439 TTCGATGTCACAGCTAACATGATGTACACCGGATGGTGGATGGTGGATGGTGGATGG 498  
 Db 119 -----TyrThrGlyVal-----ThrValSerPheLeuGly 128  
 Qy 499 ACCAGTAC-----ACATTGAGTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 549  
 Db 129 ArgArgPheGlyAspTrpArgPheSerTrpSerTyrIleLeuGlyTyrValAlaLeuLeu 148  
 Qy 550 CTCACACTAATGAGGGGTGTGATG---ATGTCATC-----GCCTGGCGGGGCG 594  
 Db 149 MetThrPhePheAlaGlyIlePheThrMetCysAlaTyrArgMetHisGlnCysArgArg 168  
 Qy 595 CTGGCA 600  
 Db 169 LeuSer 170  
 RESULT 3  
 A48300  
 N:Alternate names: lens fiber cell membrane protein MP18, lens fiber membrane intrinsic f  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 C/Accession: A48300; A34490; S05010  
 R:Gutekunst, K.A.; Rao, G.N.; Church, R.L.  
 Curr. Eye Res. 9, 955-961, 1990  
 A>Title: Molecular cloning and complete nucleotide sequence of the cDNA encoding a bovine  
 A/Reference number: A48300; MUID:91114349; PMID:2276272  
 A/Accession: A48300

A:Molecule type: mRNA  
 A:Residues: 1-173 <GUT>  
 A:Cross-references: UNIPROT:P20274; UNIPARC:UPI000012E779; GB:U04188; NID:g163294; PIDD:  
 R.Louis, C.F.; Hur, K.C.; Galvan, A.C.; Tembroek, E.M.; Javris, L.J.; Eccleston, E.D.; H  
 U. Biol. Chem. 264, 19967-19973, 1989  
 A:Title: Identification of an 18,000-Dalton protein in mammalian lens fiber cell membra  
 A:Reference number: A34490; MUID:90062105; PMID:2584203  
 A:Accession: A34490  
 A:Molecule type: protein  
 A:Residues: 1-23 <LOU>  
 A:Cross-references: UNIPARC:UPI0000178CB1  
 A:Experimental source: lens  
 R.Rao, G.N.; Gutekunst, K.A.; Church, R.L.  
 FEBS Lett. 250, 483-486, 1989  
 A:Title: Bovine lens 23, 21 and 19 kDa intrinsic membrane proteins have an identical ami  
 A:Reference number: S05010; MUID:89325619; PMID:2473922  
 A:Accession: S05010  
 A:Molecule type: protein  
 A:Residues: 1-20 <RAO>  
 A:Cross-references: UNIPARC:UPI0000178CB2  
 C:Superfamily: growth arrest-specific protein  
 C:Keywords: membrane protein  
 F:1-173/Product: lens membrane protein MP19 #status predicted <MAT>

Alignment Scores:  
 Pred. No.: 0.00669 Length: 173  
 Score: 133.00 Matches: 51  
 Percent Similarity: 39.3% Conservative: 28  
 Best Local Similarity: 25.4% Mismatches: 78  
 Query Match: 8.5% Indels: 44  
 DB: 2 Gaps: 10

US-10-537-002-7 (1-786) x A48300 (1-173)

```

QY      22  GCCTGGGCTTGGTTCATGATGGGATGGGCGGATGTCGCCACCTGCATG  81
DB      6  |||||:::|||||:::|||||:::|||||:::|||||:::|||||
      6  GlyGlyGlyLeuPheCysAlaTrpValGlyThrIleLeuValValAlaThrIleThr  25

QY      82  GACCACTGAGACCCCAAGCTTGTACAACACCCCGTAAGCTGTTTCAATCCAG  141
DB      26  |||||:::|||||:::|||||:::|||||:::|||||:::|||||
      26  AspiHisTrpMetGlnTrpArgLeu-----SerGlyAlaPheMetIleGln  40

QY      142  GGGCTGGGCGCTCCTGTCGTCGAGAGAGCTGTCGCTTCAACGAGTCCGGGCTTATTC  201
DB      41  |||||:::|||||:::|||||:::|||||:::|||||:::|||||
      41  GlyLeuTrpArgTyrCysLeuGly-----ThrIleCys-----TyrIleu  53

QY      202  ACCCTGCTGGGCTCCGACCATGCTGCAGAGCGAGTCCGAGCCCTGATGCTGATGATC  261
DB      54  |||||:::|||||:::|||||:::|||||:::|||||:::|||||
      54  GlnThrGlnSerIle---AlaTyrTrpAsnAlaThrArgAlaPheMetIleLeuSer  72

QY      262  GTCCTGGGTCATTTGGCTCGTGGATCCATCTTGGCCCTAAATGATCGCATTTGGC  321
DB      73  |||||:::|||||:::|||||:::|||||:::|||||:::|||||
      73  LeuCysAlaThrSerGlyIleIleMetGlyIleValAla-----PheAla  87

QY      322  AGCATGAGAGACTGCGCAAGCCAAAGCATGACATGACATCCGCGGATCATGTTTATGTC  381
DB      88  |||||:::|||||:::|||||:::|||||:::|||||:::|||||
      88  GlnGlnProThrPheThrArgLeuSerArgProPheSerAlaGlyIleMetPheMetAla  107

QY      382  TCAGGCTCTTGGCAATGCTGGAAGTGTGTGTTGCCAATGCTGATGCTACTTATTC  441
DB      108  |||||:::|||||:::|||||:::|||||:::|||||:::|||||
      108  Ser-----ThrPhe  111

QY      442  TGATGTCACAGTACATGATACACCGGATGGGATGATGATGATGATGATGATGATGATG  501
DB      112  |||||:::|||||:::|||||:::|||||:::|||||:::|||||
      112  ValLeuLeuAlaLeuAlaIleTyrThrGlyVal-----ThrValSerPheLeuGlyArg  129

QY      502  AGGATAC-----ACATTTGTCGCGCTCTGTTGTCGAGGCTGAGTCTGAGAGCCCTC  552
DB      130  |||||:::|||||:::|||||:::|||||:::|||||:::|||||
      130  ArgPheGlyAspTrpArgPheSerTrpSerTyrIleLeuGlyTrpValAlaLeuMet  149

QY      553  ACACATTAATTTGGGGGTGTATG---ATGTGCATC-----GCCGTCGGGCGGCTG  597
DB      150  |||||:::|||||:::|||||:::|||||:::|||||:::|||||
      150  ThrPhePheAlaGlyIlePheTyrMetCysAlaTyrArgMetCnIleGlnCysArgArgLeu  169
  
```

QY 598 GCA 600  
 DB 170 Ser 170

RESULT 4  
 G02355  
 tumor-associated membrane protein TMP - human  
 C:Species: Homo sapiens (man)  
 C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 18-Sep-1998  
 C:Accession: G02355; J05731  
 R:Ben-Porath, I.; Benvenisty, N.  
 Submitted to the EMBL Data Library, December 1995  
 A:Reference number: H01104  
 A:Accession: G02355  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-157 <BEN1>  
 A:Cross-references: UNIPARC:UPI00002F10A; EMBL:U43916; NID:g1171355; PID:g1171356  
 R:Ben-Porath, I.; Benvenisty, N.  
 Gene 183, 69-75, 1996  
 A:Title: Characterization of a tumor-associated gene, a member of a novel family of gene  
 A:Reference number: J05730; MUID:97149281; PMID:896089  
 A:Accession: J05731  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-8,'P',10-31,'N',33,'V',35-44,'T',46,'IS',49-157 <BEN2>  
 A:Cross-references: UNIPARC:UPI000178CAE; GB:U43916  
 C:Comment: This protein is involved in tumor formation and embryogenesis.  
 C:Genetics:  
 A:Gene: TMP  
 C:Superfamily: growth arrest-specific protein  
 C:Keywords: glycoprotein  
 F:2-21/Domain: transmembrane #status predicted <TM1>  
 F:67-84/Domain: transmembrane #status predicted <TM2>  
 F:95-116/Domain: transmembrane #status predicted <TM3>  
 F:134-154/Domain: transmembrane #status predicted <TM4>

Alignment Scores:  
 Pred. No.: 0.00893 Length: 157  
 Score: 131.50 Matches: 45  
 Percent Similarity: 40.2% Conservative: 29  
 Best Local Similarity: 24.5% Mismatches: 71  
 Query Match: 8.4% Indels: 39  
 DB: 2 Gaps: 7

US-10-537-002-7 (1-786) x G02355 (1-157)

```

QY      31  TTCGTGTTTCACTGATTTGGATTTGCGGCATCATGCTGCCACCTGATGACCAAGTGG  90
DB      9  |||||:::|||||:::|||||:::|||||:::|||||:::|||||
      9  TyrValValHisIleAlaThrValIleMetLeuPheValSerThrIleAlaAsnValTrp  28

QY      91  AGCACCCAAAGCTTGTACAACACCCCGTAACAGCTGTTTCACTACCAAGGCGTGG  150
DB      29  |||||:::|||||:::|||||:::|||||:::|||||:::|||||
      29  LeuValSerSerThrAlaAspAlaSerVal-----GlyLeuTrp  41

QY      151  CGCTCTGTCGCGAGAGAGCTTGGCTTCAACCGAGTCCGCGGCTACTTCAACCTGCTG  210
DB      42  |||||:::|||||:::|||||:::|||||:::|||||:::|||||
      42  LysAsnCysSerAsn-----MetGlnCysSerAspSerLeuSerTyrAla  56

QY      211  GGGCTGCAGCCATGCTGCAGGACATGCGAGCCCTGATGATGTGATGATGATGATGATG  270
DB      57  |||||:::|||||:::|||||:::|||||:::|||||:::|||||
      57  SerGlnAspAla---LeuIleThrValGlnIlePheMetIleLeuSerIleIlePheCys  75

QY      271  GCCATGGCTCTCTGCTGATTCATCTTTGCGCTGAAATGCAATCCGATTTGGAGCATGG  330
DB      76  |||||:::|||||:::|||||:::|||||:::|||||:::|||||
      76  ValIleAlaLeuLeuValPheAlaPhe-----GlnLeuPheThrMetGln  90

QY      331  GACTGTCCAAAGCCAAATGACATGACATGCTCCGGGATCATGTTTCAATGCTCAGATCTT  390
DB      91  |||||:::|||||:::|||||:::|||||:::|||||:::|||||
      91  -----LysGlyAsnArgPheMetLeuSerGlyAlaThrThrLeuValCysThrIleu  107

QY      391  TGTGCAATTTGCTGAGAGTGTGTGTTTCCAAACATGCTGTGTGATTAATCTTGTGATGTC  450
  
```

DB 108 CysIleLeuValGlyValSerIleTyr----- 116  
QY 451 ACAGTAACATGATACCGGAGTGGATGGATGGACAGCTTTCAGACCAAGTACACA 510  
DB 117 ---ThrsSerHleTyrAlaAsnArgAspGly-----ThrlGlnTyrHis 129  
QY 511 TTGTGGGGGCTCTGTGGTGGGCTGGAGGCTGACACTTAATGGGGGTGG 570  
DB 130 HlsgIlyTyrSerTyrIleuGlyTyrIleCysPheCysPheSerPheIleIleGlyAl 149  
QY 571 ATGATGTCATC 582  
DB 150 LeuTyrLeuVal 153  
RESULT 5  
S21721  
growth arrest-specific protein - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C/Accession: S21721  
R/Suter, U.; Welcher, A.A.; Oezcelik, T.; Snipes, G.J.; Kosarova, B.; Francke, U.; Billitt  
Nature 356, 241-244, 1992  
A/Title: Trembler mouse carries a point mutation in a myelin gene.  
A/Reference number: S21721; MUID:92204231; PMID:1552943  
A/Accession: S21721  
A/Status: Preliminary  
A/Molecule type: mRNA  
A/Residues: 1-160 <STU>  
A/Cross-references: UNIPROT:P16646; UNIPARC:UPI00000040F7  
C/Superfamily: growth arrest-specific protein  
C/Keywords: transmembrane protein

Alignment Scores:  
Pred. No.: 0.0228 Length: 160  
Score: 126.50 Matches: 48  
Percent Similarity: 40.6% Conservative: 25  
Best Local Similarity: 26.7% Mismatches: 58  
Query Match: 8.1% Indels: 49  
DB: 2 Gaps: 8

US-10-537-002-7 (1-786) x S21721 (1-160)

QY 61 ATCATGCTGCACCTGATGACCAAGTGG-----AGCAACCAAGACTTG 105  
DB 19 LeuPheValSerThrIleValSerGlnTyrIleuValGlyAsnGlyHisThrThrAspLeu 38  
QY 106 TACAACAACCCGCTAAGAGCTGTTTCAACTACAGAGGCTGGGCTCTGTGCGA 165  
DB 39 TrpGlnAsnGlyThrThrSerAlaLeu-----GlyAlaValGlnHisCysTyr--- 54  
QY 166 GAGAGCTGTGGCTTACCGAGTGGCGGGGCTACTTACCGCTGGGGCTGCCAGCATG 225  
DB 55 ---SerSerSerValSerGln-----Trp 61  
QY 226 CTGACGAGGAGGCGGAGCTGTATGATCGAGGATCGTCTGGGCTGCGCTCTG 285  
DB 62 LeuGlnSerValGlnAlaThrMetIleLeuSerValIlePheSerValLeuAlaLeuPhe 81  
QY 286 GTATCATCTTTGCGCTGAATGATCGCATTCGATGGCAGCATGAGAACTCTCCAAAGCC 345  
DB 82 LeuPhePheCysGlnLeu----- 87  
QY 346 AACATGACATGACCTCCGGATCATGTTGATGCTCTTGGCAAT---GCT 402  
DB 88 ---PheThrLeuThrIleGlyGlyArgPheTyrIleThrGlyPhePheGlnIleLeuAla 106  
QY 403 GGAGGTCTGTGTTTGGCAACAGCTGGTGAATCTTGATGATGCAACAGTAAATG 462  
DB 107 GlyLeuCysVal-----MetSerAlaAlaAlaIle 116  
QY 463 TACACCGGATGGGATGGATGGATGGATGAGACTTTCAGACCAAGTACATTTGGTGGCT 522

DB 117 TyrThr---ValArgHisSerGlnTyrPheValaLeuThrAspTyrSerTyrGlyPheAla 135  
QY 523 CTGTTGCGGCTGGGCTGGTGGGAGGCTTACACTTAATGGGGGTGATATGTCATC 582  
DB 136 TyrIleLeuAlaTyrValAlaPheProLeuAlaLeuSerGlyIleIleTyrValIle 155

#### RESULT 6

JC5730  
tumor-associated membrane protein precursor ; mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 09-Dec-1997 #sequence\_revision 09-Dec-1997 #text\_change 13-Nov-1998  
C/Accession: JC5730  
R/Ben-Porath, I.; Benvenisty, N.  
Gene 183, 63-75, 1996  
A/Title: Characterization of a tumor-associated gene, a member of a novel family of genes  
A/Reference number: JC5730; MUID:97149281; PMID:8956089  
A/Accession: JC5730  
A/Status: nucleic acid sequence not shown  
A/Molecule type: mRNA  
A/Residues: 1-160 <BN>  
A/Cross-references: UNIPARC:UPI000002P109; GB:U25633  
C/Comment: This protein is involved in tumor formation and in embryogenesis.  
C/Genetics:  
A/Gene: TMP  
C/Superfamily: growth arrest-specific protein  
C/Keywords: glycoprotein  
F/1-16/Domain: signal sequence #status predicted <SIG>  
F/69-87/Domain: transmembrane #status predicted <TM>  
F/93-118/Domain: transmembrane #status predicted <TM2>  
F/138-157/Domain: transmembrane #status predicted <TM3>  
F/35,43,128/Binding site: carbohydrate (Aan) (covalent) #status predicted

Alignment Scores:  
Pred. No.: 0.0582 Length: 160  
Score: 121.50 Matches: 37  
Percent Similarity: 45.8% Conservative: 23  
Best Local Similarity: 28.2% Mismatches: 50  
Query Match: 7.8% Indels: 21  
DB: 2 Gaps: 6

US-10-537-002-7 (1-786) x JC5730 (1-160)

QY 31 TTGCGTGTTCACATGATGGATGGGATTCGCGCATTCGTCACCTGATGACCAAGTGG 90  
DB 9 PheValValHisIleAlaThrAlaIleMetLeuPheValSerThrIleAlaAsnValTrp 28  
QY 91 AGCAACCAAGACTTGTATACACACCCGTAACAGCTGTTTCACTACCAAGGAGCTGTGG 150  
DB 29 MetValAlaAspTyrAlaAlaAsnAlaSerVal-----GlyLeuTrp 41  
QY 151 CGCTCTGTGTCCAGAGAGCTTGGCTTACCGAGTGGCGGGGCTACTTACCGCTGCTG 210  
DB 42 LysAsnGlyCys-----ThrGlyGly-----AsnGlyAspGlySerLeuSer---Tyr 55  
QY 211 GGGGTCGACGATGCTGACGAGTGGAGGCTGTGATGCTGTGATGCTGTGGAT 270  
DB 56 GlyAsnGlnAspAlaIleLysValValGlnAlaPheMetIleLeuSerIleLeuPheSer 75  
QY 271 GCCATTGGCTTCTGTGATTCATCTTGGCTTGAATGATCATTCGATGGCAGCATGAG 330  
DB 76 IleIleSerLeuValValPheValPhe-----GlnLeuPheThrMetGln 90  
QY 331 GACTTGGCAAAAGCCAAATGACATGACACTCCGGGATCATGTTATGTCTCAGTCTT 390  
DB 91 -----LysGlyAsnArgPhePheLeuSerGlySerThrMetLeuValCysTrpLeu 107  
QY 391 TGTGCAATTCGAGTGTGTGTTGGCAAC 423  
DB 108 CysIleLeuValGlyValSerIleTyrThrHis 118  
RESULT 7  
JN0503  
peripheral myelin protein 22 - human

[illegible]

Db LeuSerLeuPheLeuPheHeCyS-----Gln 86

Qy 346 AACATGACATGACCTCCGGGATCATGTTCAATGTCTCAGGCTTTGTCGAATT---GCT 402

Db 87 LeuPheThrLeuThrLysSerGlyGlyArgPheTyrlThrGlyIlePheGlnIleLeuAla 106

Qy 403 GGAGTGTCTGTGTTTGGCCAAACATGCTGTGACTAACTTTCGATGTGCACAGCTAACATG 462

Db 107 GlyLeuCySVal-----MetSerAlaAlaAlaIle 116

Qy 463 TACACCGGATGGGATGGATGTGACACTGTTCAAGACAGGTACACATTTGGCGGCT 522

Db 117 TyrThr---ValArgHisProGluTrpHisLeuAsnSerArgPyrSerTyrlGlyPheAla 135

Qy 523 CTGTTGCGTGGGCTGTGGCTGCAGGCCCTCACACTAATTGGGGCTGTGATGATGTGCATC 582

Db 136 TyrIleLeuAlaTrpValAlaIaPheProLeuAlaLeuSerGlyValIleTyrlValIle 155

RESULT 8

T38459

hypothetical divergent repeat-containing protein - fission yeast (Schizosaccharomyces pombe)

CtSpecies: Schizosaccharomyces pombe

CtDate: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 09-Jul-2004

CtAccession: T38459, T38380

R.Harris, D.; McDonald, S.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, February 1996

A.Reference number: Z21794

A.Accession: T38459

A.Molecule type: DNA

A.Residues: 1-1748 <HAR>

A.Cross-references: UNIPROT:Q10172, UNIPARC:UPI0000169027, EMBL:Z69368, PIDD:CAA93290.1;

R.McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.; Wood, V.

submitted to the EMBL Data Library, September 1997

A.Reference number: Z21790

A.Accession: T38380

A.Molecule type: DNA

A.Residues: 1451-1794 <MC1>

A.Cross-references: UNIPARC:UPI000016208D, EMBL:Z70691, PIDD:CAA94638.1; GSPDB:GN00066; S

CtAccession: T38380

A.Experimental source: clone c25G10

CtGenetics:

A.Gene: SPAC27F1.01c; SPDB:SPAC25G10.09c

A.Map position: 1

Alignment Scores:

Pred. No.: 0.0837 Length: 1794

Score: 118.50 Matches: 73

Percent Similarity: 37.8% Conservative: 73

Best Local Similarity: 25.1% Mismatches: 99

Query Match: 7.6% Indels: 83

Db: 2 Gaps: 15

US-10-537-002-7 (1-786) x T38459 (1-1794)

Qy 749 TCGTCTCTGTGCGGGACCTCCATCATGTATCTTCTGTTTGGTGTGACCCCAAG 690

Db 1372 SerSerAsnValHisAlaProProProProValGlnProMetAsnAlaMetProSerHis 1391

Qy 689 CCAAGTCTGGCCCTTGAAGCCTCCAGGCTTGTAGGCAACACTGTGGCCTTAGATGATA 630

Db 1392 AsnAlaValAsnAlaArgProSerAla-----ProGluArg----- 1403

Qy 629 GAAAGGGGCTTTGTAGTTTCTTCTGTGGGCCAGGCCCGGAGCGGATGCATATCATC 570

Db 1404 ArgArgSerPheGlySerValSerSerGlySerAsn-----ValSerSerIle 1419

Qy 569 ACACCCCCAATTAGTGTGAGGCT---CCAGCACCACCCAGCCAGAACAGACGGCACCA 513

Db 1420 GluArgGluThrSerThrMetProLeuLysAlaSerGlnProThrAsnProGlyAlaIle 1439

Qy 512 AATGTGTACTGTGTCTGGAACAGTCTGCACCATCCCAACCCATGCGCGTGTACATTTAGCT 453

Db 1440 SerAsnHis-----AlaProGlnValValProProAlaProMetHisAlaValAla 1456



A>Note: sequen extracted from NCBI backbone (NCBIP:106581)  
C:Comment: The predicted signal sequence may not be cleaved in some cell types.  
C:Superfamily: growth arrest-specific protein  
C:Keywords: glycoprotein; Schwann cell; transmembrane protein

Alignment Scores:

Pred. No.:	Score:	Length:
0.123	117.50	160
Percent Similarity:	41.1%	Matches: 43
Best Local Similarity:	24.6%	Conservative: 29
Query Match:	7.5%	Mismatches: 64
DB:	2	Indels: 39
Gaps:	7	

US-10-537-002-7 (1-786) x A41144 (1-160)

OY 61 ATCATTCGCGCCACCTGCATGAGCACCGAACCACCAACTTGTACACACCCGCTA 120  
Db 19 LeuPheValSerThrIleValSerGlnTrp-----LeuValGlyAsnGly--- 33  
OY 121 ACAGCTGTTTAACATAACACGAGGGGCTGTGGCGCTCCGTGTGCCAGAGAGACTTGCGCTTC 180  
Db 34 -----HisArgThrAspLeuTrpGlnAsnCysThrThrSerAlaLeuGlyAla 49  
OY 181 ACCGAGTCCGGGGCTACTTACACCTGTGGGGCTGCCACCATGCTGCAGGACGTGCGA 240  
Db 50 ValGln-----HisCysTyrSerSerSerValSerIleuTrpLeuGlnSerValGln 66  
OY 241 GCCCTGATGATGTAGGAGCATGCTGTCGGGGGCGCATGAGCTCCCTGATGCATCTTTC 300  
Db 67 AlatrMetIleuSerValIlePheSerValIleuSerIleuPheLeuPhePheCysGln 86  
OY 301 CTGAATGTCATCCGATTGTCAGACATGAGAGACTCTGCCAAAGCAAATGACACTGACC 360  
Db 87 Leu-----PheThrIleuThr 91  
OY 361 TCCGGGATCATATTCATTCCTGCAGGCTCTTTCGCAAT--GCTGAGATCTCTGTGTTT 417  
Db 92 LysGlyGlyAlaRpheryrlleThrglyValPheGlnIleuValAsglyLeuCysVal-- 110  
OY 418 GCCAACATGCTGTGATCTAATCTTGATGTCCACAGCTAATGATGACACCGGATGGGT 477  
Db 111 -----MetSerAlaAlaIleIleTyrlrHr---ValArg 120  
OY 478 GGGATGTGTGCAGACTGTTCAGACCCAGGTACACTTGTGTGGCGCTCTGTTCGTGGGCTGG 537  
Db 121 HisSerGIuTrpHlaValaaAnaMaMprTyserTylGlyPheAlaTyrlleLeuAlaTrp 140  
OY 538 GTGCTGAGGCGCTCACAATAATGGGGGTGATGATGATGATGATGATGATGATGATGATGAT 582  
Db 141 ValAlaIlePheProLeuAlaLeuLeuSerGlyIleIleTyrlleValIle 155

RESULT 10

H-transporing two-sector ATPase (EC 3.6.3.14) chain I - Methanococcus jannaschii  
G64327  
C:Species: Methanococcus jannaschii  
C:Dates: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
C:Accession: G64327  
R:Bult, C.V.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
erson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurest, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kalne, B.P.; Borodovsky, M.; Klank, H.P.; Frazer, C.M.; Smith, H.O.; Moese, C.  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.  
A:Reference number: A643300; PMID:96537999; PMID:8688087  
A:Accession: G64327  
A>Status: Preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-695 <BDU>  
A:Cross-references: UNIPROT:Q57675; UNIPARC:UPI00001381E5; GB:U67478; GB:L77117; NID:9155  
C:Genetics:  
A:Map position: REV214151-212064  
A:Start codon:-GTG  
A:Keywords: hydrolase

## Alignment Scores:

Pred. No.:	0.211	Length:	695
Score:	114.00	Matches:	54
Percent Similarity:	42.7%	Conservative:	39
Best Local Similarity:	24.8%	Mismatches:	75
Query Match:	7.3%	Indels:	50
DB:	2	Gaps:	11

US-10-537-002-7 (1-786) x G64327 (1-695)

```

QY 25 TTGGGGTTCTGTCAGTTCATGTCAGATTCGCG-----GGCATCATTCGTCGCC 72
DB 425 LeuGlyTYRLeuLeuThrLeuAlaGlyIleSerThrValIleMeGlyIleIleThrGly 444
QY 73 ACCTGCATGAGCAGCAGTGCAGACCCCAAGCTTGTACACACCCCGTAA-----123
DB 445 GlyTYRLeuGlyAspPhe---ThiTYRGluphePheGlyPheAspValThrIysThrPro 463
QY 124 ---GCTGTTTCAATCCAGGCGGCTGCGCTCCTGTCGAGAGAGCTCGCTTC 180
DB 464 LeuAlaLeuValAspProLeuGly-----GluSerTYRtyrIle 476
QY 181 ACCGATGCGCGGCTACTTCACTTCGTCGCGCTGCGCAGCCATGTCGAGCAGTGCAG 240
DB 477 AsnAsnAsnProLeuPheThr---LeuGlySerIleSerValThrAsnGlyPromet 495
QY 241 GCCCTGATGATGTCGAGGATGCTGCTGCGGTCGATTCGCTGCTGCTGATTCATCTTGGC 300
DB 496 AlaIleLeuValPheSerIlePheValGlyLeuIleIleLeuIleIleGlyLeuPhe---514
QY 301 CTGAATGATCCGATTCGTCAGATGTCGAGCATCTGCGCAAGCCAACTGACATGACGACC 360
DB 515 -----ValGlyPheIysGluAsnValIysArgGlyAspMet-----526
QY 361 TCCGGATCATGTCATTCATGTCACAGTCTTTCGTCATTCGTCGATTCGTCGTT---417
DB 527 ---GlyAspAlaPheIleAsnGlnGlyValTrpIleLeuLeuIleIleSerIlePheVal 545
QY 418 -----GCCAATGCTGTCGTCATCTTTCGATCTTCGATGTC---450
DB 546 GlyIleGlyLeuMetPheAlaGlyIleAsnThrMetIleAlaGlyIleIleGlyIle 565
QY 451 -----ACAGCTAATCATGTCACCGGCG-----ATGGTGGAGATGNG 486
DB 566 PheValIleValLeuAlaIleLeuAlaSerMetTYRlyGlyTYRlySerGlyValIleMet 585
QY 487 CAGACTGTCAGACAGGATGACATTCGTCGCGCTGTCGTCGTCGTCGTCGTCGTCGTCGA 546
DB 586 GluAlaIle-----LeuGlyAlaMetAspValThrGlyPheLeuGlyAsn 600
QY 547 GGCCTCACACTAATTCGGGCTGTGATGATGTCATGTCGCTGCGCGGCTGCGCA 600
DB 601 ValLeuSerTYRAlaArgLeuLeuAlaLeuCySLeuAlaThrGlyGlyLeuAla 618

```

## RESULT 11

JC5044

epithelial membrane protein 2 - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1997 #sequence\_revision 31-Jan-1997 #text\_change 09-Jul-2004

C:Accession: JC5044

R:Taylor, V.; Suter, U.

Gene 175, 115-120, 1996

A:Title: Epithelial membrane protein-2 and epithelial membrane protein-3: Two novel mem

A:Reference number: JC5044; MUID:97074659; PMID:8917086

A:Accession: JC5044

A:Molecule type: mRNA

A:Residues: 1-167 &lt;TAY&gt;

A:Cross-references: UNIPROT:P54851; UNIPARC:UPI000016AF4; EMBL:X94770; NID:91359880; PI

C:Comment: This protein belongs to peripheral myelin protein 22 family which are involve

C:Superfamily: growth arrest-specific protein

, Alignment Scores:

Pred. No.:	0.881	Length:	167
Score:	107.00	Matches:	47
Percent Similarity:	39.0%	Conservative:	29
Best Local Similarity:	24.1%	Mismatches:	73
Query Match:	6.9%	Indels:	46
DB:	2	Gaps:	10

US-10-537-002-7 (1-786) x JC5044 (1-167)

```

QY 25 TTGGGGTTCTGTCAGTTC-----CTGATGGGATTCGGGCGATTCGTCGCACTGC 78
DB 5 LeuAlaPheIleIleAlaPheHisIleThrSerAlaAlaLeuLeuIleAlaThrVal 24
QY 79 ATGACGACGTGAGCAGCAGCAGACCTTGTACAAACCCCGTAACTGTTTCAATAC 138
DB 25 AspAsnAlaTrpTrpValGlyAspGluPhe-----Phe 35
QY 139 CAGGCGCTGTCGCTCCTGTCGTCGAGAGAGCTGTCCTTCACTGAGTGC-----189
DB 36 AlaAspValTrpArgIleCysThrAsnAsn-----ThrAsnCysThrValIle 51
QY 190 -----CGGGGCTACTTCACTTCGTCGCGGCTGCGCAGCCATGTCGAGCAGTG 237
DB 52 AsnAspSerPheGlnGlyIysSerThr-----LeuGlnAlaPhe 64
QY 238 CGAGCCCTGATGATGTCGAGGATGCTGCTGCGGTCGATTCGCTGCTGATTCATCTTT 297
DB 65 GlnAlaThrMetIleLeuSerThrIleLeuGlyCysIleAlaPhePhe-----IlePhe 82
QY 298 GCCCTGAATGATCCGATTCGTCGAGCATGAGCATGAGCATCTGCCAAGCAATGACACTG 357
DB 83 ValLeuGlnLeuPheArgLeuGlnGlyGlu-----ArgPheValLeu 97
QY 358 ACCTCCGGATCATGTCATTCATGTCACAGTCTTTCGATTCGTCGATTCGTCGTT 417
DB 98 ThrSer---IleIleGlnLeuMetSerCysLeuGlyValMetIleAlaIleSerIleTYR 116
QY 418 GCCAATGCTGTCGTCATCTTTCGATGTCACAGCTTCAATGATGACCGGATGAGT 477
DB 117 ThrAspAlaArg---GluAspIleHisAspIysAsnAlaIysPheTYR-----131
QY 478 GGGATGTCGACAGCTTTCAGACAGGATGACATTCGTCGCGCTGTCGTCGCTGCG 537
DB 132 -----ProValThrArgGlnGlySerTYRlyGlyTYRlySerTYRlyLeuAlaTrp 147
QY 538 GTTCGTCGAGGCTCACACTAATTCGGGCTGTGATGATGTCGTCGTCGTCGTCGTCGTC 592
DB 148 ValAlaPheAlaCySThrPheIleSerGlyMetCysTYRleuIle 162

```

## RESULT 12

T45059

hypothetical protein Y398B.gg [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004

C:Accession: T45059

R:Wilson, R.; Ainscough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton, C.

raser, A.; Pulton, L.; Gardner, A.; Green, P.; Hawkins, T.; Hillier, L.; Jier, M.; Johnst

B.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rikken, L.; Roopra, A.; Saunders, D.

Nature 368, 32-38, 1994

A:Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.

A:Reference number: 843531; MUID:94150718; PMID:7906398

A:Accession: T45059

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-735 &lt;WIL&gt;

A:Cross-references: UNIPROT:O9NES7; UNIPARC:UPI0000798CB; EMBL:AL132896; NID:96434440; I

C:Experimental source: clone Y39B6B

C:Genetics:

A:Map position: 3

A:introns: 18/1; 69/1

A&gt;Note: Y39B6B.gg





Query Match: 6.6% Indels: 194  
 DB: 1 Gaps: 16  
 US-10-537-002-7 (1-786) x GYHU (1-593)

```

OY 57 GGGCATCATTTGCTGSCACCTGCATGAGCCAGTGGAGCAACCAAGACTTGTATACAGAACCC 116
    |||||
DB 95 GlyH1eH1aCySeCyS-----Pro 100
OY 117 CGT-----AACAGCTTTTCACTACAGG-----143
    |||||
DB 101 ArgGlyPheH1aCySeSerAlaAspGlyArgSerCySPheGlnArgSerGlyAsnAsnSer 120
OY 144 GCTGTCGGCCTCTGTCGCGAGAGCTCTGCTTACACGAGCGCGGCTACCTTAC 203
    |||||
DB 121 ValGlyAlaIleGlnCyPro-AspSerGln-----PheGlnCySPProAspPhe-SerT 138
OY 204 CCTGC-----TGGGCTGCACGACCATGTCGACGAGCTGCGAGC 242
    |||||
DB 138 hrCyCyValMetValAspGlySerTyrGlyCyCySerProMetProGln-AlaSerCyS 157
OY 243 CCTGATGATCTTAGGCAT-----CGTCTTGGG 269
    |||||
DB 158 CySPGlnAspArgValH1aCySeCySProH1aGlyAlaPheCyAspLeuValH1aThrArg 177
OY 270 TGGCATTTGGCCTCTGTCATTCAT-----CTCAGG 293
    |||||
DB 178 CysIleThrProThrGlyThrH1aProLeuAlaIleGlySerProAlaGlnArgThrAsn 197
OY 294 -----CTTGCCTGGAATGCATCCGATTTGGAGCAT 326
    |||||
DB 198 ArgAlaValAlaLeuSerSerSerValMetCySPProAspAlaArgSerArgCySPProAsp 217
OY 327 GGA-----GACTCTGCCAAGCCAAATGAC 353
    |||||
DB 218 GlySerThrCySPGlyLeuProSerGlyGlyTyrGlyCySPProMetProAsnAla 237
OY 354 ACTGACCTCCGGAGCATGTTCAATGT-----CTCAGG 386
    |||||
DB 238 ThrCySPSerAspH1aIleuH1aCySPProGlnAspThrValCyAspLeuIleGln 257
OY 387 TCTTTGTGC-----395
    |||||
DB 258 SerIysCySPLeuSerIleGlyAsnAlaThrThrAspLeuLeuThrIleuProAlaH1a 277
OY 395 -----395
    |||||
DB 278 ThrValGlyAspValIysCyAspMetGluValSerCySPProAspGlyTyrThrCySPCyS 297
OY 396 -----ATTGCTGAGAGTGTCTGT-----GTTTGCCAAACATGCT 428
    |||||
DB 298 ArgLeuGlnSerGlyAlaTyrGlyCySPProPheThrGlnAlaValCyS-----314
OY 429 GGTGACTAATTCTGATGTCCACAGCTAACATGTACACCGGATGGGTGGAGTGTCA 488
    |||||
DB 314 -----CTTGCCTGGAATGCATCCGATTTGGAGCAT 326
    |||||
OY 489 GACTGTTCAGACAGGTACAC-----ATTGTCGGGCTCTGTTCTGCGG 533
    |||||
DB 315 ---CySPGlnAspH1aIleH1aCySeCySPProAlaGlyPheThrCyAspThrGlnIlyGly 333
OY 534 CTGGGTCCCTGGAGCCTCAC---ACTAATTGG-----563
    |||||
DB 334 ThrCySPGlnGlnGlyProH1aIleGlnValProThrMetGluValProAlaH1aIleuSer 353
OY 564 -----GGGTGATGATGTGCATGCGCTGCGG 590
    |||||
DB 354 LeuProAspProGlnAlaLeuIlyAspArgAspValProCySPAspAsnValSerSerCySPPro 373
OY 591 GGGCTGGACCAAGAAACCAACTACAAAGCCTTTCTTATCATGCTCAGGCGACACAG 650
    |||||
DB 374 SerSerAspThrCySPGlnLeuThrSerGlyGlnTyrGlyCySPCySPProIleProGln 393
  
```

OY 651 -----TGTTCCTACAGCCTGG 668  
 DB 394 AlaValCySPSerAspH1aIleGlnH1aCySPProGlnIlyTyr 408

RESULT 14  
 G53610  
 ntp protein - Enterococcus hirae  
 C:Species: Enterococcus hirae  
 C:Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 18-Sep-1998  
 C:Accession: G53610  
 R:Takebe, K.; Kakimura, S.; Yamato, I.; Konishi, K.; Igataashi, K.; Kakimura, Y.  
 J. Biol. Chem. 269, 11037-11044, 1994  
 A:Title: Sequencing and characterization of the ntp gene cluster for vacuolar-type Na(+)-  
 A:Reference number: A53610; MUID:94209269; PMID:8157629  
 A:Accession: G53610  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-448 <TAK>  
 A:Cross-references: UNIPARC:UPI0000178EC9; GB:D17462  
 C:Superfamily: Na+-ATP synthase chain J

Alignment Scores:  
 Pred. No.: 1.89 Length: 448  
 Score: 102.50 Matches: 47  
 Percent Similarity: 44.1% Conservative: 28  
 Best Local Similarity: 27.6% Mismatches: 64  
 Query Match: 6.6% Indels: 31  
 DB: 2 Gaps: 10

US-10-537-002-7 (1-786) x G53610 (1-448)

```

OY 139 CAGGGGCTGTGGCGCTCTGTGTCGAGAGACTGTGCTTACCGAGTGGCGGGCTTAC 198
    |||||
DB 157 LysGlyIleThrPheSerIlePheH1aValaSerSerPhe-----CysAsnAlaGly 174
OY 199 TTCACCTGTCTGGG-----CTGCCACCATGTCTGACGAGCTGCGAGCCTGATGATC 252
    |||||
DB 175 PheAspLeuLeuGlyAspSerLeuLeuAlaAspGlnThrAsnValTyrLeuIleMetVal 194
OY 253 GTAAGC-----ATCGTCTGGTGCCATTTGGCCTCCCTGATTCATCTTGGCCTGAAA 306
    |||||
DB 195 ValSerAlaLeuIleIleAlaGlyLeuGlyPheIleValITrpaArgAspIleLeuSer 214
OY 307 TGCAATCCGATTTGGAGCATGAGAGACTGTGCCAAGCCAAAGCAATGACATGACCTCC--- 363
    |||||
DB 215 TyrH1aGlyValIlySerIleThrLeuH1aSerIysValAlaLeuSerValThrAlaLeu 234
OY 364 -----GGGATCATGTTCAATGTC-----TCAAGTCTTTGTGCAATT 399
    |||||
DB 225 LeuLeuIleGlyGlyPheIleLeuPheLeuIleThrGluArgAsnGlyLeuThrLeuVal 254
OY 400 GCTGAGAGTGTGTGTTGGCCAAACATGCTGCTGACTACTTCTGATGTCC-----450
    |||||
DB 255 LysGly-----ThrPheThrGluArgLeuAlaAsnThrPheMetSerValThrPro 272
OY 451 ---ACAGCTAACATGTACAC-----GGCATGGGAGGATGGTGCAG 489
    |||||
DB 273 ArgThrAlaGlyTyrTyrSerIleAspTyrLeuGlnMetSerH1aIleGlyLeuIleLeu 292
OY 490 ACTGTTTCAGACAGCATACATTTGGTGGCGCTCTGTTGAGCTGGGTGCTGAGAGC 549
    |||||
DB 293 ThrMetPheLeuMetTyrIleGlyGlyThrSer-----GlySerThrAlaGlyGly 309
OY 550 CTC---ACACTAATTTGGGGGTGTGATGATG 576
    |||||
DB 310 LeuIlyThrThrThrLeuGlyIleLeuLeu 319
  
```

RESULT 15  
 AD3550  
 NMD(P) trianhydrogenase (AB-specific) (EC 1.6.1.2) - Brucella melitensis (strain 16M)  
 C:Species: Brucella melitensis  
 C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 31-Dec-2004  
 C:Accession: AD3550





Score:	101.00	Matches:	45
Best Local Similarity:	39.3%	Conservative:	30
Best Local Similarity:	23.6%	Mismatches:	64
Query Match:	6.5%	Indels:	52
DB:	2	Gaps:	8





```

Db 1701 -----ProGlyTTPrelnglyProValGlyGlnGlnAlaSerValAlaSerGlyCys 815
QY 507 -----CACATTGGTGGGCTGTCTGTCGAGGCTGGGCTGC 542
Db 816 AlaGlyAlaSerProGlyProHlaGlyThrCysThrAlaSerProGlyAlaSerPheArg 835
QY 543 -----TGAGAGCCTACACATAATTGGGGGTGAT---GATCGATGAGCCG--- 587
Db 836 CysIleCysHlaArgGlyTyrThrGlyProPheCysAlaSerGlnAlaSerValAlaSerCysAla 855
QY 588 CCGGGGAGCCT-----GGACACAGAGAAACCACTACAAAGCCGTTTCTTATCA 635
Db 856 ProAlaProCysLeuHlaGlyGlySerCysGlnAlaSerGlyValGlySerPheSerCysSer 875
QY 636 TGGCTC-----AGGCCACAGTGT-----TGCCTTACAGGC 665
Db 876 CysLeuAlaSerGlyPheAlaGlyProArgCysValAlaArgValAlaSerGlyCysLeuSerSer 895

RESULT 22
S45306
notch 3 protein - mouse
C/Species: Mus musculus (house mouse)
C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C/Accession: S45306
R/Lardelli, M.; Dahlstrand, J.; Lendahl, U.
Mech. Dev. 46, 123-136, 1994
A/Title: The novel Notch homologue mouse Notch 3 lacks specific epidermal growth factor-
A/Reference number: S45306; MUID:95001556; PMID:7918097
A/Accession: S45306
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-2318 <LNR>
A/Cross-references: UNIPROT:061982; UNIPARC:UPI00002930C; EMBL:X74760; NID:9483580; PID:
C/Suprafamily: notch protein; ankyrin repeat homology; EGF homology
F/163-195/Domain: EGF homology <EGF1>
F/474-505/Domain: EGF homology <EGF>
F/854-885/Domain: EGF homology <EGF2>
F/1839-1871/Domain: ankyrin repeat homology <AN1>
F/1872-1904/Domain: ankyrin repeat homology <AN2>
F/1906-1938/Domain: ankyrin repeat homology <AN3>
F/1939-1971/Domain: ankyrin repeat homology <AN4>
F/1972-2004/Domain: ankyrin repeat homology <AN5>

Alignment Scores:
Pred. No.: 2.4 Length: 2318
Score: 100.50 Matches: 64
Percent Similarity: 28.4% Conservative: 19
Best Local Similarity: 21.9% Mismatches: 64
Query Match: 6.4% Indels: 145
DB: 2 Gaps: 17

US-10-537-002-7 (1-786) x S45306 (1-2318)
QY 702 GTTGAACCCAAAGCCAGTCTGCTGCTGTAAGCTCCAGGCTTGTAGCAACACTGTGGCC 643
Db 1167 LeuGlyProSerLeuAlaSerGlyValGlnCys----- 1177
QY 642 TAGGAGCATGATGAAGAACGCTTT-----GTAGTTGGTTTCTTGTGTCGACAGCC 592
Db 1178 -----LeuHlaSerGlyThrCysValAlaSerValGlyGlyPheAlaArgCysAlaCys 1194
QY 591 CCGGAGGAGGATGACATCATCACACCCCAATTAAGTGAAGCCCTCCAGCAACCAAGCC 532
Db 1195 ProProGlyTyrThrGlyLeuHla-----CysGlnAlaAlaSerLeu----- 1208
QY 531 CACGAACAGAGCCGACCAATGTGATGCTGTGAACAGTGTGACCATCCACCCAT 472
Db 1209 -----GlnCysArgProGly-----AlaCysHla 1216
QY 471 GCGGATGATCATGTAGTGTGTGACATCCAGAAATTAGTCAACAGATTTGGCAACAC 412
Db 1217 AlaAla----- 1218

```

```

QY 411 AGACATCCAGCAATTTGCAAAAGACTGAGACATGAACATGATCCGAGAGCTAG--- 355
Db 1219 ---HlaThrAlaArgAlaSerCysLeuGln-----AspProGlyValHlaPhe 1231
QY 354 -----TGTCAATTTGGCTTTTGGCAGAGTCTTCATGC----- 323
Db 1232 ArgCysValCysHlaProGlyPhe---ThrGlyProArgCysGlnAlaLeuSerProC 1251
QY 322 -----TGCATTCGAGATGATTCAGGGCAAGATGATGACAGAGGCC 277
Db 1251 YeGlySerGlnProCysGln----- 1257
QY 276 AATGGACACCCAGAGAGAGCTTACAGATCATCAGGCTC----- 239
Db 1258 -----HlaGlyValGlnCysArgHlaSerLeuGlyAlaGlyGlyLeuThrPheThrC 1276
QY 238 -----GCATTCGCTTCAGCATGC----- 221
Db 1276 YeHlaCysValProProHlaThrGlyLeuArgCysGlnAlaValAlaArgSerCysArgG 1296
QY 220 ---CTGGCAGCCCGCAGAGGTAAGTAAGTACCCGCGCATCGTGAAGCCGAGACTCTTCG 163
Db 1296 LuleGlnCysProValGlyIleProCys-----GlnGlnThrAlaArgG 1311
QY 162 GACACAGAGGCGCCACAGCCCTGGTGAATGAACAGCTGTACGGGGT-----TGT- 110
Db 1311 LProArgCysAlaCysProProGly-----LeuSerGlyProSerCysAla 1326
QY 109 -----TGTACAGTCTTGGTGGCTCCACTGTCATGCAATGCA 76
Db 1326 rgValSerArgAlaSerProSerArgAlaThrAlaAlaSerCysAlaSerAlaProCysL 1346
QY 75 GGTGGACGATGATGATGCCCGCATCCATCCATCGTGA--- 38
Db 1346 euHlaGlyGlySerCysLeuProValGlnSerValProPheArgCysValAlaAlaP 1366
QY 37 -----CCAGGACCCCAAGCCCT 20
Db 1366 roGlyTTPrelGlyProArgCysGlnThrPro 1376

RESULT 23
PC2061
genome polypeptide N2 - hepatitis C virus
N/Contains: envelope protein E1; nonstructural protein E2/NS1
C/Species: hepatitis C virus
C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004
C/Accession: PC2061
R/Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.
Biochem. Biophys. Res. Commun. 199, 1474-1481, 1994
A/Title: Identification of the third major genotype of hepatitis C virus in France..
A/Reference number: PC2060; MUID:94197744; PMID:8147893
A/Accession: PC2061
A/Molecule type: mRNA
A/Residues: 1-411 <LNR>
A/Cross-references: UNIPROT:081813; UNIPARC:UPI00000900B; GB:U12355; NID:9410169; PID:
C/Suprafamily: hepatitis C virus genome polypeptide; glycoprotein; nonstructural protein;
C/Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;
F/192-383/Product: envelope protein E1 #status predicted <SP>
F/384-411/Product: nonstructural protein E2/NS1 #status predicted <NP>
F/196,209,234,305,325/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 3.04 Length: 411
Score: 100.00 Matches: 57
Percent Similarity: 36.8% Conservative: 34
Best Local Similarity: 23.1% Mismatches: 102
Query Match: 6.4% Indels: 55
DB: 2 Gaps: 11

US-10-537-002-7 (1-786) x PC2061 (1-411)
QY 26 TGGGATTCGTTGTTCACTGATTTGGATTTGGGATTCAGGAGCATTCGTCGCACTGATGAGCC 85

```

```

Db      93  TPrLaGLYtrPLeu-----LeuSerProArgGlySer-----ArgProSerTrpGly 108
QY      86  AGTGAAGACCCCAAGACTTGTACAAACACCCGGTAACACTGTTTCACTACCCAGGGC 145
Db      109  ProAnaMPProArg-----ArgArgSerLeuArgAn 118
QY      146  TGTGGCGCTCTGTGTCCAGAG-AGCTGTGCTTCACAGAGTCCGGGGCTACTTACCC 204
Db      119  LeuGLyLysValILleAspThrLeuThrCysGlyPheAlaAspLeuMetCylYrIlePro 138
QY      205  CTGCTGGGGCTGCCA-----GCCATCTGCAGGCACTGGAGCCCTG 246
Db      139  LeuValGLyAlaProValGLyGLyValAlaArgAlaLeuAlaHISGLYValArgAlaLeu 158
QY      247  ATGATC---GTAGGATCGTCCGGGTGCATT---GGCTCTGTGTATCCATCTTTGCC 300
Db      159  GluAspGLyILleAsnPheAlaThrGLyAsnLeuProGLyCysSerPheSerIlePheLeu 178
QY      301  CTG-----AAATGCATCCGATTTGGCAGACATGGAGACTGTCCCAAGCCAAACATG 351
Db      179  LeuAlaLeuPheSerCysLeu-----IleHisProAlaAlaSerLeuGLUTrp 194
QY      352  ACACTGACCTCCGGGATCATGTTCAATGCTTTCAGACTTTTGCATTTCCAGATGCT 411
Db      195  ArgAnThrSerGly---LeuTyrlleuThrAsnAspCysSerAsnSerSerIleVal 213
QY      412  GTGTTTGCACATGCTGTGTACTTCTC-----TGATGTCCACAGCT 456
Db      214  TyrGLuAlaAspAspValILleuHISThrProGLyCysValProCysValGLInaAspGly 233
QY      457  AACATGTACACCGGATGAGTGTGTCAGACTGTTCAGACCAAGTAC----- 507
Db      234  AsnThrSerThrCysTrpThrProValThrProThrValAlaValArgTrpValGLyAla 253
QY      508  -----ACATTGGTGGC 519
Db      254  ThrThrAlaSerILleArgSerHISValAspLeuLeuValGLyAlaAlaThrMetCysSer 273
QY      520  GCTCTGTTCGTGGGCTGGGTGCTGAGGCTCACAACCTAATTTGGGGGTGTATGATGTC 579
Db      274  AlaLeuTyValGLyAspMetCysGLyAlaValPheLeuValGLyGLInaIleAspThrPhe 293
QY      580  ATCGCTGCGGGGCTGGCAGACCAACCAACTCAAGCCGTTTCTTATCATGCC 639
Db      294  ArgProArgArgHISGLInThrValGLInThrCysAsnCysSerLeuTyProGLyHISLeu 313
QY      640  TCAGGCCACAGTGTGCCTAC 660
Db      314  SerGLyHISArgMetAlaTrp 320

RESULT 24
S41288
Genome polypeptide - hepatitis C virus (fragment)
NContains: core protein; envelope protein; NS1 protein
CSpecies: hepatitis C virus
CDate: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 31-Dec-2004
CAccession: S41288
R:Seelig, R.
submitted to the EMBL Data Library, December 1993
AReference number: S41288
AAccession: S41288
A.Molecule type: genomic RNA
A.Residues: 1-492 <SEED>
A.Cross-references: UNIPROT:068870; UNIPARC:UPI0000178522; EMBL:X76918
A.Keywords: capsid protein; core protein; envelope protein; nonstructural protein; polyP
F1-191/RProduct: core protein #status predicted <CON>
F192-372/RProduct: envelope protein #status predicted <ENV>
F373-492/RProduct: NS1 protein (fragment) #status predicted <NS1>

Alignment Scores:
Pred. No.: 3 Length: 492
Score: 100.00 Matches: 56
Percent Similarity: 36.8% Conservative: 35

```

```

Best Local Similarity: 22.7% Mismatches: 102
Query Match: 6.4% Indels: 55
Db: 2 Gaps: 11

US-10-537-002-7 (1-786) x S41288 (1-492)

QY      26  TGGGGTGTGGTTCACATGATGGAGATGGAGCTATGCTGSCACTGATGAGACC 85
Db      93  TPrLaGLYtrPLeu-----LeuSerProArgGlySer-----ArgProSerTrpGly 108
QY      86  AGTGAAGACCCCAAGACTTGTACAAACACCCGGTAACACTGTTTCACTACCCAGGGC 145
Db      109  ProAnaMPProArg-----ArgArgSerLeuArgAn 118
QY      146  TGTGGCGCTCTGTGTCCAGAG-AGCTGTGCTTCACAGAGTCCGGGGCTACTTACCC 204
Db      119  LeuGLyLysValILleAspThrLeuThrCysGlyPheAlaAspLeuMetCylYrIlePro 138
QY      205  CTGCTGGGGCTGCCA-----GCCATCTGCAGGCACTGGAGCCCTG 246
Db      139  LeuValGLyAlaProValGLyGLyValAlaArgAlaLeuAlaHISGLYValArgAlaLeu 158
QY      247  ATGATC---GTAGGATCGTCCGGGTGCATT---GGCTCTGTGTATCCATCTTTGCC 300
Db      159  GluAspGLyILleAsnPheAlaThrGLyAsnLeuProGLyCysSerPheSerIlePheLeu 178
QY      301  CTG-----AAATGCATCCGATTTGGCAGACATGGAGACTGTCCCAAGCCAAACATG 351
Db      179  LeuAlaLeuPheSerCysLeu-----IleHisProAlaAlaSerLeuGLUTrp 194
QY      352  ACACTGACCTCCGGGATCATGTTCAATGCTTTCAGACTTTTGCATTTCCAGATGCT 411
Db      195  ArgAnThrSerGly---LeuTyrlleuThrAsnAspCysSerAsnSerSerIleVal 213
QY      412  GTGTTTGCACATGCTGTGTACTTCTC-----TGATGTCCACAGCT 456
Db      214  TyrGLuAlaAspAspValILleuHISThrProGLyCysValProCysValGLInaAspGly 233
QY      457  AACATGTACACCGGATGAGTGTGTCAGACTGTTCAGACCAAGTAC----- 507
Db      234  AsnThrSerThrCysTrpThrProValThrProThrValAlaValArgTrpValGLyAla 253
QY      508  -----ACATTGGTGGC 519
Db      254  ThrThrAlaSerILleArgSerHISValAspLeuLeuValGLyAlaAlaThrMetCysSer 273
QY      520  GCTCTGTTCGTGGGCTGGGTGCTGAGGCTCACAACCTAATTTGGGGGTGTATGATGTC 579
Db      274  AlaLeuTyValGLyAspMetCysGLyAlaValPheLeuValGLyGLInaIleAspThrPhe 293
QY      580  ATCGCTGCGGGGCTGGCAGACCAACCAACTCAAGCCGTTTCTTATCATGCC 639
Db      294  ArgProArgArgHISGLInThrValGLInThrCysAsnCysSerLeuTyProGLyHISLeu 313
QY      640  TCAGGCCACAGTGTGCCTAC 660
Db      314  SerGLyHISArgMetAlaTrp 320

RESULT 25
T01437
Hypochemical protein R34001.1 - human
CSpecies: Homo sapiens (man)
CDate: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
CAccession: T01437
R:Hammerlin, J.E.; McCreedy, P.M.; Skowronski, E.; Adamson, A.W.; Burkhat-Schultz, K.; G
P.; Christensen, M.; Georgescu, A.; Avila, J.; Liu, S.; Bruce, R.; Quan, G.; Krommiller,
M.; Coefield, J.; Amico-Keller, G.; Lucas, S.; Duarte, S.; Olsen, A.O.; Carrano, A.V.
submitted to the EMBL Data Library, February 1998
A.Description: Sequence analysis of a 1 Mb region in 19q13.1.
AReference number: Z14330
AAccession: T01437
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA

```

A/Residues: 1-1007 <LIM>  
A/Cross-references: UNIPROT:P043379, UNIPARC:UPI000006CC3, EMBL:AC004444, NID:92887496;  
C/Genetics:  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
A/Protein: 82/2, 113/1, 155/1, 177/3, 218/2, 243/3, 332/2, 398/2, 416/1, 487/3, 513/2, 5  
Alignment Scores:  
Pred. No.: 2.82 Length: 1007  
Score: 100.00 Matches: 45  
Percent Similarity: 31.3% Conservative: 12  
Best Local Similarity: 24.7% Mismatches: 45  
Query Match: 6.4% Indels: 81  
DB: 2 Gaps: 11

US-10-537-002-7 (1-786) x T01437 (1-1007)

```

OY 75 CTGATGACGACAGT-----GAGCACCAGACTGTACAA 110
DB 842 LeuHISGlySerAlaPheArgProSerLeuProAlaProGluSerProGlyLeuProAla 861
OY 111 CAACCCCGTAACAGCTGTTTCACTACAGAGGGCTGT-----GGGCTCTGTGTCCGAG 166
DB 862 HLeProSerAsnPro-----GlnLeuProGluAlaArgProGly11aProGlyLeuThr 879
OY 167 AGAGCTCTGCTTCAACGAGCCGGGCTTACTTCACTGCTGGGGCTGCAGCAGCATGC 226
DB 880 AlaSerLeuLeuGluProThrSerGlyTTPGlyThrSerCysThrGlyCysArgProPro 899
OY 227 TGCAGGACAG-----TGCAGGCCCTGATGATCGTAGGCATGCTCTG 268
DB 900 SerLysLysProSerThrPheThrValCys-----TTP 910
OY 269 GAGCCATGAGCTCTGCTGTATTCATCT-----TTGCCCTGA 304
DB 911 SerProValAlaArgTTPThrProGlySerSerArgHISGlyLeuSerTTPSerProPro 930
OY 305 AATGATCCGATGAGCATGAGAG-----ACTGCGCAAG 343
DB 931 SerCysGlySerThrAlaSerTTPArgLeuAlaTTPTrpGlyLeuValTTPProGln 950
OY 344 CCACATGACACTGACCTCCGGATCAGTTCATGTCAGAGCTTTGTCAAATTGCTG 403
DB 951 ProArg----- 952
OY 404 GAGTGTCTGTGTGGCA-----ACATGCTGTGACTA 436
DB 953 -----LeuCysArgAlaGlnAspProArgProHISArgArgCysThrProTTP----- 968
OY 437 ACTTGTGATGTTCACAGCTAACATGATACAGCGGATG-----GTGGGATGG 484
DB 969 -----ProAlaGlnThrCysArgProCysTTPAsnThrThrArgSerCysTTP 984
OY 485 TGCAGA 490
DB 985 CysArg 986

```

RESULT 26

S74569  
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 2 - *Synechocystis* sp. (strain PCC 68  
N/Alternate names: hypothetical protein ndhB  
C/Species: *Synechocystis* sp.  
A/Variety: PCC 6803  
C/Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
A/Accession: S74569, A39366  
R/Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asami, Y., Miyajima, N.,  
O. K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda  
DNA Res. 3, 109-136, 1996  
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*  
A/Reference number: S74322; MUID:97061201; PMID:8905231  
A/Accession: S74569  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-521 <KAN>

A/Cross-references: UNIPROT:P2714; UNIPARC:UPI000013065D; EMBL:D90900; GB:AB001339; NID:  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
R/Ogawa, T.  
Proc. Natl. Acad. Sci. U.S.A. 88, 4275-4279, 1991  
A/Title: A gene homologous to the subunit-2 gene of NADH dehydrogenase is essential to tr  
A/Reference number: A39366; MUID:91239550; PMID:1903537  
A/Accession: A39366  
A/Molecule type: DNA  
A/Residues: 1-411, 'R', 413-521 <OG>  
A/Cross-references: UNIPARC:UPI00000BAED0, GB:D90288; NID:9217096; PDB:BAI4330.1; PDB:  
C/Genetics:  
A/Gene: ndhB  
C/Superfamily: NADH dehydrogenase (ubiquinone) chain 2  
C/Keywords: membrane-associated complex; NAD; oxidoreductase

Alignment Scores:  
Pred. No.: 3.6 Length: 521  
Score: 99.00 Matches: 69  
Percent Similarity: 34.6% Conservative: 33  
Best Local Similarity: 23.4% Mismatches: 75  
Query Match: 6.3% Indels: 118  
DB: 2 Gaps: 17

US-10-537-002-7 (1-786) x S74569 (1-521)

```

OY 4 GCCGTGAGTCCCTGTCAAGGGCTTGGGTTTCGTTGATGAT-----GGG 51
DB 205 AlaAspThrValGlyGlnSerLeuValAlaLeuValPheVal11aLeuGly 224
OY 52 ATTGGGGCATCATGCTGTCCACCTGATGAGCAGTGGAGACCCAGACTGTACAC 111
DB 225 ILeaAlaPheLysIleSerAlaValProPheHISGlnTTP-----ThProAspValArgL 243
OY 112 AACCCGTAACAGCTGTTTCACTACAGAGGGCTGTG----- 150
DB 244 GlySerProThrProValAlaPheLeuSerValGlySerLysAlaLeuGlyPheAla 263
OY 151 -----CGTCTCTGTGTCGAGAGAGCTCTGCTTACAC-----GAGTGGCGGGGCTAC 198
DB 264 ValAlaIleArgLeuLeuValThrAlaPheGlyGlyIleHISAspGlnTTPHISValIle 283
OY 199 TTCACCTGTGTGGGGCTCCAGCCATG-----CTGCAGGAGTGGGAGCCCTG----- 246
DB 264 PheThrAlaLeuAlaValLeuSerMetValLeuGlnValAlaLeuAlaGlnThr 303
OY 247 ----- 258
DB 304 SerMetLysArgMetLeuAlaTyrSerSerIleGlyGlnAlaGlyPheValMetIleGly 323
OY 259 ATCGTCCCTGGTGGCC-----ATTGGCCCTC 282
DB 324 LeuValAlaGlySerGlnAspGlyTyrAlaSerMetValPheTyrMetLeuIleTyrLeu 343
OY 283 CTGTATCCATCTTTCCTCCGTAATGATC-----CGCATTTGGCAGCATG 327
DB 344 PheMetLeuLeuGlyAlaPheSerCysIleIleLeuPheMetLeuArgThrGlySerAsp 363
OY 328 GAGACTTGCACAAAGCCACATG-----ACACTGACTCCGGATC----- 369
DB 364 GlnIleSerAspTyrAlaGlyLeuTyrHISLysAspProLeuLeuThrLeuIleLeuSer 383
OY 370 -----ATGTTATCTGTCTCAGAGCTTGTGTGATGCTGAGTGTCTGTGTGCAAC 423
DB 384 IleCysLeuLeuSerLeuGlyGlyIleProProLeuAlaGly-----PheHISGlyLys 401
OY 424 ATGCTGTGATCACTTGTGTGATGTCCACAGCTAACATGTACACCGGATGGTGGGATG 483
DB 402 IleTyrIle-----PheTTP----- 406
OY 484 GTGCAGACTTTCAGACAGGATACATTTGTGCGGCTGTGTTGGGCTGTGGCT 543
DB 407 -----AlaGlyTTPGlnSer 411

```





A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-714 <KUR>  
A:Cross-references: UNIPROT:Q8UG67; UNIPARC:UPI00000D1A7A; GB:AE007869; PIRID:AAK86977.1;  
C:Genetics:  
A:Gene: AGR\_C\_2169  
A:Map position: circular chromosome  
C:Superfamily: H(+) -translocating inorganic pyrophosphatase

Alignment Scores:

	Pred. No.:	4.23	Length:	714
Score:	98.00	Matches:	60	
Percent Similarity:	41.2%	Conservative:	45	
Best Local Similarity:	23.5%	Mismatches:	80	
Query Match:	6.3%	Indels:	70	
DB:	2	Gaps:	15	

US-10-537-002-7 (1-786) x H97502 (1-714)

```
QY 34 GTGTTTCACTGATGGATGCGGCGCATATGCTGCACCTGCATGACGACGATGAGC 93
DB 8 lIeVallIleuCyegIyValIeuserValIyTAlaVal-----TrrThr 23
QY 94 ACCCAA-----GACTGTACAACAAACCCGTAACAGCTGTTTCAACTAC 138
DB 24 ThrLysSerValIeuaSpAlaSpGInglYArGMeTArgGluIleAlaGlyTyr 43
QY 139 CAGGGGCTGTGGCGCTCTGTGTCCGAGAGACCTGCTGCATCCGAGCCGGGCTAC 198
DB 44 -----IleArgGluGlyAlaGlnAlaTyrIleuThrArgGlnTyr 56
QY 199 TTCACC-----CTGCTGGGCGCTGCCAGCC-----ATGCTGCAGGA 234
DB 57 LeuThrIleAlaIleValIeGlyLeuIleValAlaValLeuAlaTrrPtyrIleuSerAla 76
QY 235 GTGCAGCCCTG--ATGATCGTAGGCATGTCCTG-----GGTCCATTGCG 279
DB 77 lIeAlaAlaIleGlyPheValIleGlyAlaValIeuserGlyValAlaGlyPheValGly 96
QY 280 CTCCTGATTCATCATCTTGGCC---CTGAATGATCCGATTCGAGCATGAGACACT 336
DB 97 MetHisValSerValArgAlaAsnIeuaGlnThrAlaAlaSerHisSerIeugly 116
QY 337 GCCAAGCCACATGACATGACTCCGGGATCATG-----TTCATTGTCAGGCTT 390
DB 117 AlaGlyLeuAspIleAlaPheLysSerGlyAlaIleThrGlyMetLeuValAlaGlyLeu 136
QY 391 TGTGCAATTGCTGAGTGTCTGTGTTGCCAACATGCTGTGACT----- 435
DB 137 ---AlaLeuLeuGlyValSerIleTyrTyrPheValIleuThrSerValIeuglyHisPro 155
QY 436 -----AACTTCGATGTCACAGT 456
DB 156 ProGlySerAlaGlyAlaValIleAspAlaLeuValSerLeuGlyPheGlyAlaSerIeule 175
QY 457 AACATGATCACCGGATGGTGGATGTGTGAGACTGTTTCAGACGAGTACATTTGGT 516
DB 176 SerIlePheAlaArgLeuGlyGlyIle-----PheThrIysGly 189
QY 517 GCGGCTCTG-----TTCTGTGGCTGGGTCCGCTGAGGCTC----- 552
DB 190 AlaAspValGlyGlyAspLeuValGlyLysValGlnAlaGlyIleProGluAspAspPro 209
QY 553 -----ACGTAATTGGGGGTGTGATGATGTGATGCC---TGGCGGGGCGCTG 597
DB 210 ArgAsnProAlaThrIleAlaAspAsnValGlyAspAsnValGlyAspCysAlaGlyMet 229
QY 598 GCACCAAGAAACCAACTACACAAAGCCGTTTCTTATCATGCTCA 642
DB 230 AlaAlaAspLeuPheGluThrTyrThrAlaValSerValValAlaThr 244
```

nascent polypeptide-associated complex alpha chain, muscle splice form gp220 - mouse  
N:Alternate names: alpha-NAC protein  
C:Species: Mus musculus (house mouse)  
C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
R:Yotcov, W.V.; St-Arnaud, R.  
Genes Dev. 10, 1763-1772, 1996  
A:Title: Differential splicing-in of a proline-rich exon converts alphaNAC into a muscle-  
A:Reference number: Z20889; MUID:96312450; PMID:8698236  
A:Accession: T30826  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2187 <YOT>  
A:Cross-references: UNIPROT:P70670; UNIPARC:UPI00000275BE; EMBL:U48363; NID:gl66668; PIR  
C:Genetics:  
A:Gene: Naca  
A:Map position: 10  
A:Introns: 24/1; 1996/1; 2050/3; 2099/3; 2142/3; 2183/3  
A>Note: differential splicing converts alphaNAC into a tissue-specific DNA-binding active  
C:Keywords: alternative splicing; DNA binding; transcription factor

Alignment Scores:

	Pred. No.:	3.86	Length:	2187
Score:	98.00	Matches:	76	
Percent Similarity:	34.9%	Conservative:	30	
Best Local Similarity:	25.0%	Mismatches:	111	
Query Match:	6.3%	Indels:	87	
DB:	2	Gaps:	16	

US-10-537-002-7 (1-786) x T30826 (1-2187)

```
QY 734 GCACCTCATGTAATATCTTGTGTTTGGTGTGACCCAAAGCCATGCGCTTG 675
DB 1501 AlaProProSerGlu---ArgValThrThrValProProGluThrAlaThrProGln 1519
QY 674 AAGCTCAGGCTTGTAGGCAACACTGTGCGCTGAGCATGATTAAGAAAGGCTTTGAG 615
DB 1520 LysAlaSerGlyThrThrAlaSerLysValProValPro---AlaGluThrGlnGluVal 1538
QY 614 TTGCTTTTCTGCGGCCAGGCGCCCGGAGGAGACATCATCATGACCCCATTTAGT 555
DB 1539 AlaValSerSerArgGluThrPro-----ValThrProAlaVal--- 1551
QY 554 GTGAGGCTTCAGCAGCCAGCCACGAAC-----AGAGCCGACCAATGTGTACTG 501
DB 1552 -----ProProValLysAsnProSerSerHisLysLysThrSerLysThrIleGluLeu 1569
QY 500 GTCTGAACAGTCTGACCATCCACCCATGCCGGGTGATCATGTTAGCTGTGACATCCAG 441
DB 1570 LysGluAlaProAlaThrLeuProProSerProThrLysSerProLysIleProSerSer 1589
QY 440 AAGTTA----- 435
DB 1590 LysAlaValaProAlaGlyThrSerAlaProLysGluPheProAlaSerProSerIleLysPro 1609
QY 434 GTCAACAGCATGTTGGCAACAGACATCCAGCAATGTGACAAAGACTGACATG 375
DB 1610 ValThrThrSerLeuAlaGlnThrAlaProProSerLysGlnLysAlaProSerThr--- 1628
QY 374 AACATGATCCGGAG-----GTCACTGTCAATGTTGGCTTTGGCAGAGTCC 330
DB 1629 ---ThrIleProLysGluAsnLeuAlaAlaProAlaValIleuProValSerSerLysSer 1647
QY 329 TCCATGCTGCCAATGCGGATGATTCATTCAGGGCAAGATGATPACGAGGCGCAATG--- 273
DB 1648 ProAlaAlaProAlaArgAlaSerAlaSerLeuSerProAlaThrAlaAlaProGlnThr 1667
QY 272 GCACCAAGAGATGCTTACATC----- 249
DB 1668 AlaProLysGluAlaThrThrIleProSerCysLysLysAlaAlaAlaThrGluThrPro 1687
QY 248 ATCAGGGCTCGCATGCTGTGACGATGGCTGACAGCCCGACAGG----- 204
```



```

Db 1668 IlegluthrSerThralAProserLeuGlulAProlyGluThrSerdlurhSer 1707
Qy 203 -----GTAAAGTACCCCGGACCTGGTGAAGCCGAGCTCTTGGACA 159
Db 1708 ValSerlyValLeuMet-SerSerProlyLeuValSerSerlyS- 1724
Qy 158 CAGAGCCGACAGCCCGCTGGT-----AGTGAACAAAG-----CTGTACG 117
Db 1725 ArgAlaserThrLeuProAlathrThrLeuProSerLeuLyGluAlaserValLeuSe 1744
Qy 116 GGGTGTGTACAACTCTTGG-----TGCTC 90
Db 1744 rProthrAlathrSerSerlyLeuApsSerhIleSerProValSerAlaCySe 1764
Qy 89 CACTGTCCATGACAGTGGACGCAATGATGCCGCAATCCCAATGATGAACCAAC 30
Db 1764 ThrGlyThrThr-----ThrProGlnAlaserGluLyLeuProSerly 1779
Qy 29 CCAAGCCCT 20
Db 1779 GlyGlyPro 1782

RESULT 32
T09059
notch4 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C/Accession: T09059
R/Rosen, L.; Mbehatia, G.; Qiu, S.; Ahearn, M.E.; Dankers, C.; Laeky, S.; Loretz, C.; Sc
submitted to the EMBL Data Library, October 1997
A/Description: Sequence of the mouse major histocompatibility locus class III region.
A/Reference number: Z16543
A/Accession: T09059
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1964 <ROM>
A/Cross-references: UNIPROT:P31695; UNIPARC:UPI000016C7F1; EMBL:AF030001; NID:92564945;
C/Genetics:
A/Gene: notch4
A/Map position: 17
A/Intons: 22/1; 49/2; 148/1; 264/1; 305/1; 384/1; 436/1; 501/1; 539/1; 577/1; 618/1; 67
1679/3; 1729/1; 1761/3
C/Superfamily: notch protein, ankyrin repeat homology; EGF homology
C/Keywords: receptor; signal transduction
P/514-545/Domain: EGF homology <EGF>

Alignment Scores:
Pred. No.: 4.27 Length: 1964
Score: 97.50 Matches: 60
Percent Similarity: 31.84 Conservative: 17
Best Local Similarity: 24.84 Mismatches: 75
Query Match: 6.24 Indels: 91
Db: 2 Gaps: 16

US-10-537-002-7 (1-786) x T09059 (1-1964)
Qy 763 GATAGATTTGATCTCTCTCTG-----TGCGGCACTCCATGATATCT 716
Db 1004 ArgGluCyLeuApsArgProCyhIleProserGlyThrAlaCyhIleSerleuAla 1023
Qy 715 ACTGTGTTTGGTGTGACCCAAAGCCAGTGTGAGCCCTTGAAGCCCTTGAGG 656
Db 1024 AenAlaPheLyCyS-----GlnCyLeuPro-----GlyhIleThrGly 1036
Qy 655 CAACCTGGCGCTTGAAGCATGAT-----AAGAAAGCCTTTGTAGT----- 614
Db 1037 GlnApsCyGluValGlnUeApsLeuCyGlnSerGlnProCySerhAnGlyCySer 1056
Qy 613 ---TGTTTCTTCTGTGTCAGCCCGGACGAGCCAGTGC----- 578
Db 1057 CyGluIleThrThrGlyProProProGlyhPheThrCyhIleCyAProlyGlyhGlu 1076
Qy 577 -----ACATATCATCAACCCCAATTGATGAGGCTTCAAGCA----- 539

```

```

Db 1077 GlyProthrCySerSerhIleValAlaLeuSerCyGlyIlehIleCyhIleAnGlyCy 1096
Qy 538 -----CCAGCCCAAGAACAGCCGCAACAAATGTGTACTGCTGACAGT-- 490
Db 1097 LeuCyLeuProSerProlyhProlySerProProLeuCyhIleCyS-LeuSerGlyPh 1116
Qy 489 -----CTGACACATCCCAACCCGCTGTACATGTTAGCTGAGCA 446
Db 1116 eGlyGlyProApsCyLeuThrProProAlaProProGly-----CyGlyPr 1132
Qy 445 TCCAGATTTAGTACACAGCATTTGGCAACACAGACCTCCAGCAATTCAGCAAGAC 386
Db 1132 oProSerProCyS-----LeuIleAnGlyThrCyhIleGluTh 1145
Qy 385 CTGAGACATGACATGATCCCGAGGTGACGTGCA-TGTTGGCTTTGGACAGTCTCC 327
Db 1145 r---ProGlyLeuGlyAenProGlyhPheGlnCyhIleThrCyS-----ProPr 1159
Qy 326 A----- 326
Db 1159 ApsSerProGlyProApsCyGlnApsProGlyhAlaSerGlyCyGlnGlyApsGlyG 1179
Qy 325 -----TGCTGCCAATGCGGATGATTTCAAGGCAAGATGATACCAAGGSCAATG 273
Db 1179 yApsGlyThrCyApsAlaGlyCyS---Serly-----ProGlyGlyApsTr 1194
Qy 272 GCACCCAGACAGATGCTTACGATATGACAGGCTGCACTGCTGAGCATGGCTGGACG 213
Db 1194 PAspGlyGlyApsCyS-----SerLeuGlyAlaProApsProThrLyGlyCy 1210
Qy 212 CCA 209
Db 1210 sPro 1211

RESULT 33
T14271
Doc4 protein, stress-induced - mouse
N/Alternate names: odz protein homolog
C/Species: Mus musculus (house mouse)
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
A/Reference number: Z17951; MUID:98315054; PMID:9649432
A/Accession: T14271
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-2825 <MAN>
A/Cross-references: UNIPROT:O70465; UNIPARC:UPI00000294B0; EMBL:AF059485; NID:93170614;
C/Genetics:
A/Gene: Doc4

Alignment Scores:
Pred. No.: 4.15 Length: 2825
Score: 97.50 Matches: 57
Percent Similarity: 29.74 Conservative: 9
Best Local Similarity: 25.74 Mismatches: 69
Query Match: 6.24 Indels: 87
Db: 2 Gaps: 16

US-10-537-002-7 (1-786) x T14271 (1-2825)
Qy 70 GCACTGTGATGACCACTGAGACCCAAAGCTTGTACAAACCCCGTACAGCTGT 129
Db 766 AlathrCyLeuApsArgGlnCySerdlYhIleGlyhPhe-----LeuProApsThrGlyLeu 784
Qy 130 TTCAATCAACAGGAGCTGTGCGCTCC-----TGTCCTCGAGAG 168
Db 785 CyAenCyApsProSerTrThrGlyhIleApsCySerIleGluIleCyhIleAlaAps 804
Qy 169 AGCTGTGCTTCAACCAAGTGC-----CGGGGCTACTTCAACCTGC-----TGCGGCTG 216

```

```
Db      805  CysGlyGlyHisGlyValCysValGlyGly-ThrCysArgCysGlyAspGlyTyrMetC 824
Qy      217  CAAAGCATGC---TGCAGGCACTGCAGAGCCCTGATGATGTAGGCAT-----CGTC 264
Db      824  ValAlaCysAspArgGlnArgAlaCysHis-ProArgCysAlaGlnHisGlyThrCysValG 844
Qy      265  CTGGG---TGCATATGGGCTCCGGT-----ATGCATCTTTGGC 300
Db      844  spGlyLysCysGluCysThrProGlyTyrAsnGlyGlnHisCysThrTrileGlyGlyCysP 864
Qy      301  CTGAATATGCATCCGATTTGGAGCATGAGAACTCTGCCAAGCCAAACATGACACTGACC 360
Db      864  ro-----GlyLeuGlyAsnGly----- 869
Qy      361  TCCGGGATATGTTGATTTGTCTCAAGTCTTTGTGCAATGTGCGAGTGTCTGTGTTGCC 420
Db      870  -----AsnGlyArgCysThrLeuAspLeu---AsnGlyTyrHis---CysValCysG 885
Qy      421  AACATGCTGTGATCTAACTTTCTGAGATGTCACAGCTAACATGTACACCGGATGGGAG 480
Db      885  In----- 889
Qy      481  ATGTGTCAGACTGTTCACAGACGTACACATTTGGTCGGCTCTGTTGCTGGGCTGGGTC 540
Db      889  xgGlyThrGlyCysAspThrSerMetGluThrGlyCysGlyAsp----- 903
Qy      541  GCTGAGAGGCTCACTAACTTAATGGGGGTGTGATGATGTGATGCGCTGGCGGGGCTGGCA 600
Db      904  -----GlyLysAspAsn-----AspGlyAspGly- 911
Qy      601  CCAAGAAACCAACTACAAAGCCGTTTCTTATCATGCTCTGAGCCAGCAGATGTGGCTAC 660
Db      912  -----LeuValAspCysMetAspProAspCysGlyLeuG 923
Qy      661  AA 662
Db      923  In 923

RESULT 34
A36324
growth arrest-specific protein - mouse
C:Species: Mus musculus (house mouse)
C>Date: 25-Jan-1991 #sequence_revision 25-Jan-1991 #text_change 09-Jul-2004
C:Accession: A36324
R:Maniatis, G.; Ruato, M.E.; Del Sal, G.; Philipson, L.; Schneider, C.
Mol. Cell. Biol. 10, 2924-2930, 1990
A:Title: A growth arrest-specific (gas) gene codes for a membrane protein.
A:Reference number: A36324; WUID:90258882; PMID:1692961
A:Accession: A36324
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-144 <MAN>
A:Cross-references: UNIPROT:P16646; UNIPARC:UPI00001456E3; GB:M32240
C:Superfamily: growth arrest-specific protein
C:Keywords: transmembrane protein

Alignment Scores:
Pred. No.:      6.4      Length:      144
Score:          96.50    Matches:      44
Percent Similarity: 39.2%  Conservative: 21
Best Local Similarity: 26.5%  Mismatches: 51
Query Match:    6.2%      Indels:      50
DB:             2        Gaps:        8

US-10-537-002-7 (1-786) x A36324 (1-144)
Qy      61  ATCATTTGTCACCTGCATGAGCAGTGG-----AGCACCAGACTTG 105
Db      19  LeupheValSerThrLeuValSerGlnTrpLeuValGlyAsnGlyHisThrThrAspLeu 38
Qy      106  TACAAACACCCCGTAACAGCTGTTTCAACTACAGAGGCGCTGTGGCTCTCTGTCCGA 165
Db      106  TACAAACACCCCGTAACAGCTGTTTCAACTACAGAGGCGCTGTGGCTCTCTGTCCGA 165
```

```
Db      39  TrpGlnAsnCysThrThrSerAlaLeu-----GlyAlaValGlnHisCysTyr--- 54
Qy      166  GAGAGCTGTGGCTTACCCAGAGTGCAGGGGCTACTTACCCCTGTGGGCTGCAGCATG 225
Db      55  ---SerSerSerValSerGln-----Trp 61
Qy      226  CTGCAGGAGTGCAGAGCCCTGATGATGCTAGACATCGTCTGGGTGCATTTGGCTCTGTG 285
Db      62  LeuGlnSerValGlnAlaThrMetIleLeuSerValIlePheSerValLeuAlaLeuPhe 81
Qy      286  GTATCATCTTTGGCCCTGAATATGATCCGCATTTGGCAGACATGAGAGACTCTGCCAAACC 345
Db      82  LeupheCysGlnLeu----- 87
Qy      346  AACATGACATGCATCTCCGGGATCATGTTGATGCTCTGAGGCTTTGTGCATTT---GCT 402
Db      88  ---PheThrLeuThrLysGlyGlyAlaArgPheTyrIleThrGlyPhePheGlnIleLeuAla 106
Qy      403  GGAGTGTCTGTGTTTGCACAACTGCTGTGACTAACTTCTGATGTCCACAGTAACATG 462
Db      107  GlyLeuGlyVal-----MetSerAlaAlaAlaIle 116
Qy      463  TACACCGGATGAGGTGGAGATGTGCAGACTGTTCAGACAGGTACATTTGTGTGGCT 522
Db      117  TyrThr---ValArgHisSerGlnTrpHisValAsnThrAspTyrSerTyrIlePheAla 135
Qy      523  CTGTTCGTGGGCTGGG 538
Db      136  ThrSerTrpProGly 140

RESULT 35
A35524
high-affinity branched-chain amino acid transport system permease protein 14vH [imported]
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: A35524
R:DelVecchio, V.G.; Kaprat, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, I.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Haglund, S.; O'Callaghan, D.; Letessac Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: A35524; PMID:11756688
A:Accession: A35524
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-292 <KUN>
A:Cross-references: UNIPROT:Q8YDQ4; UNIPARC:UPI000058368; GB:AE008918; PDB:1AAL53362.1;
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME110121
A:Map position: 11

Alignment Scores:
Pred. No.:      6.03      Length:      292
Score:          96.50    Matches:      37
Percent Similarity: 43.8%  Conservative: 20
Best Local Similarity: 28.5%  Mismatches: 38
Query Match:    6.2%      Indels:      35
DB:             2        Gaps:        7

US-10-537-002-7 (1-786) x A35524 (1-292)
Qy      193  GGCCTACTTCAACCTGCTGGGCTGCAGCCATGCTG-----CAGCAGTGCAGACCCCTG 246
Db      19  GlyAlaTyrAlaIleLeuGlyValCysAlaIlePheThrTyrArgLeuValAlaVal 38
Qy      247  ATGATCGTAGGCATGCTCTGGGTGTCATTTGGC-----CTCTGTATCATCTTTGGC 300
Db      39  AsnPheThrGlyAlaAlaIleGlyAlaIleGlyThrPheIleMetValAlaLeuPheGly 58
Qy      301  CTGAATGATCCGATTTGGAGCATGAGAGACTCTGCCAAGCCAAACATGACACTGACC 360
Db      59  -----AlaGlyValProLeuVal 64
```

Db 626 rghlsalgluthtsergylalvalalproalldylglyleuthrargtyrleuproiles 646

Qy 383 CAGGCTTTGTCGAATGCTGAGTGTCTGTGTTGCCAACAAGCTGTGACTACTTCT 442

Db 646 erglyvaliserxerValalalaleuSerProtyrValalenuytrhrllethrlyaspC 666

Qy 443 GGATGTCACAGCTAAACATGTACACCGGATGAGTGTGGG-----ATGATGACGA 490

Db 666 yslenProilleuaspMetGluThrGlyAsnllleglyalatyryalvalleuValaspG 686

Qy 491 CTGTTACAGACACGATGACATTTGTCGCGGCTCTGTTCCGTGGCTGG 537

Db 686 lnrthGlyAsnMetalathrArgleuArgAlalalalValProGlyTyr 701

RESULT 37

JC5732

tumor-associated membrane protein XMP - human

C:Species: Homo sapiens (man)

C:Date: 09-Dec-1997 #sequence\_revision 09-Dec-1997 #text\_change 18-Sep-1998

C:Accession: JC5732

R:Ben-Porath, I.; Benvenisty, N.

Gene 183, 69-75, 1996

A:Title: Characterization of a tumor-associated gene, a member of a novel family of genes

A:Reference number: JC5730; MUID:97149281; PMID:8996089

A:Accession: JC5732

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-167 <BEN>

A:Cross-references: UNIPARC:UPI000178CAD; GB:U52100

C:Comment: This protein is involved in tumor formation and is embryogenesic.

C:Genetics:

A:Gene: XMP

C:Superfamily: growth arrest-specific protein

C:Keywords: glycoprotein

F:2-23/Domain: transmembrane #status predicted <TM1>

F:68-84/Domain: transmembrane #status predicted <TM2>

F:95-117/Domain: transmembrane #status predicted <TM3>

F:143-163/Domain: transmembrane #status predicted <TM4>

F:144,47,52/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:

Pred. No.:	6 94	Length:	167
Score:	96.00	Matches:	48
Percent Similarity:	38.5%	Conservative:	27
Best Local Similarity:	24.6%	Mismatches:	74
Query Match:	6.2%	Indels:	46
DB:	2	Gaps:	10

US-10-537-002-7 (1-786) x JC5732 (1-167)

Qy 25 TTGGGGTTCGTCGTTTCA-----CTGATTTGGATTCGGGCAATCATGTGTCGACCTGC 78

Db 5 leuAlapheillelleAlapheillelethrserAlalaleuSeupheillelathVal 24

Qy 79 ATGACACGTGAGACACCCAGACCTTGACAAACACCCGTAACAGCTGTTTCACTAC 138

Db 25 AsphenalatrptyrValGlyaspGluPhe-----Phe 35

Qy 139 CAGGGCGTGTGGCGGCTCGTGTGCGAGAGAGCTGTGGCTTACCCAGTGC----- 189

Db 36 AlapspValITPArilleySerThrAsnAn-----ThrAnCySerhValille 51

Qy 190 -----CGGGCTACTTCACTTCCTGTCGGGGCTGCCAGCCAGCTGACGACGATG 237

Db 52 AsnapserPheGlnGluTyrSerThr-----LeuGlnAlalVal 64

Qy 238 CGAGCCCTGATGATGTGAGGACATCGTCTGGGGTGCACATTTGACCTCTGGATTCATCTTT 297

Db 65 GlnlathrMetlileleuSerlleleuCyCySyllalalPhePhe-----llePhe 82

Qy 298 GCCCGAATGATCGGATTTGGACGATGAGGACATCTGCCAAGCAACATGACACTG 357

Db 83 ValleunlneupheArgleuGlyngllyGluPro-----PheValleu 97

QY 358 ACCTCCGGGATCATTTTCATTTGCTCAGGCTTTTGTCGAAATTCCTGGAGTGTCTGTGTT 417

Db 98 ThrSer---IleIleGlnLeuMetSerCysLeuCyValMetIleAlaIleSerIleYr 116

QY 418 GCACAATGCTGTGATCACTTCTGTGATGTCACAGCTTAACATGATACACCGGATGGGT 477

Db 117 ThrSapaAaGy---GluAspIleHisAspYlsAsnAlaIySphenYr----- 131

QY 478 GGGATGTGCAGACTTGTCAACACAGGTACATTTGTGTGGCTCTGTTGCTGGCTGG 537

Db 132 -----ProValThrArgGluGlySerYrGlyIYrSerYrIleProGlyVal 147

QY 538 GTGCTGGAGGCGCTCACTAATTTGGGGGTGAGTGAATGATGTCATC 582

Db 148 GlYAlaPheAlaCySrhPheIleSerGlyMetCetylLeuIle 162

RESULT 38

A38346

ultra-high-sulfur keratin 1 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 28-Jun-1991 #sequence\_revision 28-Jun-1991 #text\_change 09-Jul-2004

C:Accession: A38346

R:Wood, L.; Miller, M.; Hatzenbuehler, N.; Vogeli, G.

J. Biol. Chem. 265, 21375-21380, 1990

A:Title: Serine-rich ultra high sulfur protein gene expression in murine hair and skin c

A:Reference number: A38346; MUID:9106560; PMID:2250030

A:Accession: A38346

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-230 <WOO>

A:Cross-references: UNIPROT:064507; UNIPARC:UPI00002103E; GB:M37759; NID:g200961; PIDN

C:Superfamily: ultra-high-sulfur keratin

Alignment Scores:

Pred. No.: 6.76 Length: 230

Score: 96.00 Matches: 54

Percent Similarity: 26.2% Conservative: 9

Best Local Similarity: 22.5% Mismatches: 68

Query Match: 6.2% Indels: 110

DB: 2 Gaps: 12

US-10-537-002-7 (1-786) x A38346 (1-230)

QY 12 TGCCCTGAGGGCTTGGGGGTCGNGTTTCACTGATTTGGGATTTGGGATCATTTGCTGC 71

Db 24 CysGlySerGlyCysGlyGlyCysGlySerHencyGlyGlyCysGlySerSerCysCys 43

QY 72 CACCTGATGAGCAGGTGA-----GCACCAAGACTTTGTATACAAACCCGTAACAGCTG 127

Db 44 Lys-----ProValCysCysCysValProValCysCysValProValCysSer 60

QY 128 TTTTCAACTACAGGGGCTGTGGGCTCTCTGTGTCGAGAGAGAGCTTGGCTTACCGAGT 187

Db 61 CysSerSerCysGlyGlyCysGlySerSerCysGlyGlyCysGlySerCysGlySerSer 80

QY 188 GCCGGG-----GCT 196

Db 81 CysGlyGlyCysGlySerSerCysCysValProValCysCysValProValCysSer 100

QY 197 ACTTACCCCTGCTGGGGCTGCCAGCCATGCTGCAGCAG-----TGCAGACCTTGA 247

Db 101 CysSerSerCysGlyGlyCysGlySerProCysCysGlnSerSerCysCysValPro--- 119

QY 248 TGATGTAGGATCTCTCTGGGTGTCATTTGGCTCTCTGTATCCATCTTTGCCCTGAAT 307

Db 120 -----CysCysSerSer-GlyCys----- 125

QY 308 GCATCCGATTGGCAGCATGAGAGACTTGCACCAAGCCAAATGACATGACTTCCGGGA 367

Db 126 -----GlySerSerCysCysGlnSer----- 133

QY 368 TCAGTTCAATGTCTCAGGCTTTTNGC-----AATTCGAGAGTGTCTGTGTTG 418

[illegible]

[illegible]

```

Db      278 -----ValPheValAlaLeuIleValLeuGlnLeuSerSeriLyLeuAsnLeu 294
Oy      559 ATTGCG-----GGTGTATGATGTGCATC-----582
           ::::|
Db      295 LeuGlyAlaAsnGlnHlaLeuAlaThrAlaLeuTrgIlyValLeuMetIleValValMet 314
Oy      583 GCTTCGCCGGGGCTTCGACCAACAAGAAACCATTACAAGCCGTTTCTTAATCATGCTCTCA 642
           |||||
Db      315 AlaAlaArgIlyLeuPheSer-----SeriYrPheAlaSer 326

RESULT 41
T45918
hypothetical protein F5K20_80 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
CjDate: 04-Feb-2000 #sequence_revision 04-Feb-2000 #ext_change 05-Oct-2004
CjAccession: T45918
R:Monfort, A.; Casasuberta, B.; Pujadomenech, P.; Mewes, H.W.; Lemcke, K.; May
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23017
A:Accession: T45918
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-361 <MON>
A:Cross-references: UNIPROT:Q9M348; UNIPARC:UPI00009E7G3; EMBL:ALJ32960
A:Experimental source: cultivar Columbia; BAC clone F5K20
A:Map position: 3
A:Introns: 93/2; 143/1; 197/3; 279/2; 321/3
A>Note: F5K20_80
CjSuperfamily: AAA protein

Alignment Scores:
Pred. No.:          6.51          Length:        361
Percent:            96.00         Matches:       59
Score Similarity:   36.98         Conservative: 45
Best Local Similarity: 20.98      Mismatches:   91
Query Match:        6.28          Indels:       87
DB:                  2            Gaps:          15

US-10-537-002-7 (1-786) x T45918 (1-361)
Oy      21 GGGCTTGGGGTTCGCGGTTTCACTGATGGGATTCGGGATCATTTGTCGCCATGCAT 80
           |||||
Db      27 GlyArgIlyValValyGlnPheArgSeriTrp-----36
Oy      81 GGACCAATGGAGAACCCGAAGACTTATACAACACCCCGTAACAGCTGTTTCACTACCA 140
           |||||
Db      37 -----PheProTrpLeuIle-----ProCySerPheValAlaIle 47
Oy      141 GGGGCTGTGGGCGCTCTG-----TGTCCGAGAAGAGCTTGGGC--177
           |||||
Db      48 AsnValAlaValPheValIleThrMetTrpValAsnAmcCyProIylAs-SerGlyAs 67
Oy      178 ----TTTACCCAGATGCCGGGGCTACTTACCC-----CTGCT 209
           |||||
Db      67 pCySerPheAlaAspPheLeuGlyArgPheSerPheGlnAsnThrArgGlyAsnProLeuIle 87
Oy      210 GGGGCTGCGACGACGANG---CTGCAGACAGTGGGAGCGCGATGATGATGTAAGCATGCTCT 266
           |||||
Db      87 uGlyProSerSerLeuThrLeuGlnThrMetCetylValLeuAspValIlyValValValIly 107
Oy      267 G-----GATGCCATTGAGCT 281
           |||||
Db      107 sGlyAspGlyIlyTrpArgLeuLeuSerCyAsnTrpLeuHlaGlyValValIleAsile 127
Oy      282 CTTGATTCATCATCTTGGCCCTGAAAATGATCCGCGATTTGGCACAGATGAGAAGATCTGCCCA 341
           |||||
Db      127 uLeuMetAsnMetLeuThrLeuLeuPheIleGlyIle-----139
Oy      342 AGCCAATGATGACATGACATCCGAGATCATGCTTATGCTTGTGCAATTCG 401
           |||||
Db      140 ---ArgMetGlyLeuValGlyIleGlyLeuLeuTrpIleIleSeriGly-----PheGl 155

```

```

QY 402 TGGAGTGTCTGTGTTGGCAACATGCTGTGACTTAATTCTGATGTCCAGCTAACAT 461
DB 155 YGLYSERILEULESERALALEUPHEULURSERSNLIESERVALGIALASERGLYAL 175
QY 462 GTACACCGGC--ATGGGTGGATGGTGCAGACTGTTCAGACCAGTACACATTGGTGC 518
DB 175 AVALPHEGLYLEULLEUGLYGLYMETLEU-----SE 185
QY 519 GGCTGTGTTCTGGGCTGGGCTC-----GCTGAGGCTTCACACTAATTGGGGGTGTGAT 572
DB 185 TGLINLEPHELEIAENRTPHRLIETYSERASNDYSVALITRILILEVALTHLEUVA 205
QY 573 GATCGTACGCCCTGCGCGG-----GGCTGGACACCAAGAAACCACTAACAAAG 623
DB 205 ILEULEVALALAVALEUENGLYLEULIYALILEUPROGLYVALASPHEALAH 225
QY 624 CGTTTCTTATCATGCTTCAGGCCACAGTGTTCCTAC-----AACCTT----- 666
DB 225 AILEGLYGLYPHEALATMRIGLYPHELEULENGLYPHEVALLEULEULIETARGPROHIE 245
QY 667 -GGAGGCTTCAAGGCCAGACACTGGCTTGGGTCCAAACACCAAAACAGAAATATGCA 725
DB 245 TGLYTRIPLEANGINARGASNGLYPROGLYALALASPHEALHARGPHELYSILETYGL 265
QY 726 TGA 729
DB 265 NGLY 266

```

## RESULT 42

PC2060

genome polyprotein N1 - hepatitis C virus

NCContains: envelope protein E1; nonstructural protein E2/NS1

C/Species: hepatitis C virus

C/Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 31-Dec-2004

C/Accession: PC2060

R.Li, J.S.; Vitvitski, L.; Tong, S.P.; Trepo, C.

Biochem. Biophys. Res. Commun. 199, 1474-1481, 1994

A/Title: Identification of the third major genotype of hepatitis C virus in France.

A/Reference number: PC2060; MUID:94197744; PMID:8147893

A/Accession: PC2060

A/Molecule type: mRNA

A/Residues: 1-411 &lt;Lit&gt;

A/Cross-references: UNIPROT:O81489; UNIPARC:UPI0000178524

C/Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;

F192-389/Product: envelope protein E1 #status predicted &lt;SPB&gt;

F192-411/Product: nonstructural protein E2/NS1 #status predicted &lt;NPB&gt;

F196,209,234,305,325/Binding site: carbohydrate (Asn) (covalent) #status predicted

## Alignment Scores:

Pred. No.:	6.44	Length:	411
Score:	96.00	Matches:	56
Percent Similarity:	36.78	Conservative:	36
Best Local Similarity:	22.38	Mismatches:	97
Query Match:	6.28	Indels:	63
DB:	2	Gaps:	11

US-10-537-002-7 (1-786) x PC2060 (1-411)

```

QY 26 TGGGTTGCTGTTTCACTGATGGATTGCGGGATCATTTGTCGACCTGATGACC 85
DB 93 TTPALAGLYTPRLEU-----LEUSERPROARGGLYSER-----ARGPROSERTRIPGLY 108
QY 86 AGTGAAGACACCCAAAGACTTGTCAACAACACCCGTAAACAGCTTTTCAACTCCAGGGGC 145
DB 109 PROAENASPPIAARY-----ARGARGSERARGASN 118
QY 146 TGTGCGCTCTGCTGTGTCGAGAG-AGCTGTGCTTCACGAGTGCAGGGGCTACTTCAACC 204
DB 119 LEUGLYLVSVALIILEASPHRLIETHRICYSGLYPHEALASPHEUMECGLYTRILEPRO 138
QY 205 CTGCTGGGGCTGCGCA-----GCCATGCTCAGAGCAGGCTGAGCCCTG 246

```

```

DB 139 LEUVALGIALAALPROVALIGLYVALALAAARGALALEUALHIEGLYVALARGALAEU 158
QY 247 ATGATC---GTAGGATGTCCTGGGTGCCATT---GGCCCTCGGTATTCATCTTGGCC 300
DB 159 GLUAPRGLYILEAPRHEALATHRGLYASNEUPROGLYCYSSERPHESTERILEPHELEU 178
QY 301 CTG-----AAATGATCCGCAATTGGCAGACTGAGAGACTGTGCCAAAGCCAAACATG 351
DB 179 LEUALALEUPHESERYALEU-----ILEANPROALALASERLEUENLURTP 194
QY 352 AACTGACCTCCGGGATGATGTTCAATGTCACAGGCTTTGTGCATTTGCTGAGT--- 408
DB 195 ARGASNTRISERGLY---LEUTYVALLEUHRASMPERYSSERASNSERISLEVAL 213
QY 408 ----- 408
DB 214 TYRGLUALASPGLUALIILEUENHISTHRPROGLYCYVALPROCYVALGLNAPRGLY 233
QY 409 ----- 447
DB 234 ASNILESERMETCYSTRTPHRSERVALPHRPROTHRVALALVALYENHIEVALGLYALA 253
QY 448 TCCACAGCTAACATGTACACCGGATGGGTGGATGTGTCAGACTGTTCAGACAGTAC 507
DB 254 THRTHRILASERILEARGSERHIEVALASPHEULEUVALIGLYALALA----- 269
QY 508 ACATTGTGGGGCTGTGCTGTGGGCTGGGCTGGGCTGGGCTGACACTTAATTGGGGT 567
DB 270 THRMECYSSERIALALEUVALGLYASPYALCYSGLYLALAVALEUVALGLYALN 289
QY 568 GTGATGATGTCATCGCTCGCGGGGCTGCGACACCAAGAAACCACTAACAAAGCGTT 627
DB 290 ALAPRHETHRPHARGPROARGARHISGLINTHRVALGINTHRVALCYASNCYSSERLEUTYR 309
QY 628 TCTTATCATGCTTCAGGCCACAGTGTTCCTAC 660
DB 310 PROGLYHISLEUSERGLYHISARGMETALATRP 320

```

## RESULT 43

B87658

NAD(P) transhydrogenase, beta subunit [imported] - Caulobacter crescentus

C/Species: Caulobacter crescentus

C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 31-Dec-2004

C/Accession: B87658

R.Nietman, W.C.; Paldblyum, T.V.; Paulsen, I.T.; Neilson, K.E.; Eisen, J.; Heidelberg, J.I.

B.; Laub, M.T.; DeBoy, P.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolome

n, J.; Brmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A/Title: Complete Genome Sequence of Caulobacter crescentus.

A/Reference number: A87249; MUID:21173698; PMID:11259647

A/Accession: B87658

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-472 &lt;STO&gt;

A/Cross-references: UNIPROT:O9A3A1; UNIPARC:UPI00000C7A19; GB:AE005673; NID:g13424997; PJ

A/Genetics: C:3303

C/Superfamily: NAD(P) transhydrogenase, beta subunit; NAD(P)+ transhydrogenase (B-specific)

## Alignment Scores:

Pred. No.:	6.37	Length:	472
Score:	96.00	Matches:	64
Percent Similarity:	35.68	Conservative:	47
Best Local Similarity:	20.58	Mismatches:	109
Query Match:	6.28	Indels:	92
DB:	2	Gaps:	13

US-10-537-002-7 (1-786) x B87658 (1-472)

```

QY 7 GTGACTGCTGTACAGGCTTGGGCTTCGCTTCA----- 42
DB 18 ILEUALALEUARGGLYLEUSERSERPROVALTHRISERGLNTHRGLYASNARGASNGLY 37

```





QY 676 AAGCC-----AGCAGCTGGCTTGGTCCAC 702  
 DB 491 SerAlaGlnLeuSerLeuAbaArgSerAlaArgAlaAlaGlyPheGlyGlySer 510  
 QY 703 ACCAAAAAC-----AAGAAGATATACGATGAGGTGCCGACCA 741  
 DB 511 ThrAapLeuLeuProGlyLysThrMetAapAapLeuGlyIleArgThr 526

RESULT 45  
 AF2632  
 Conserved hypothetical protein Atu0457 [imported] - Agrobacterium tumefaciens (strain C5  
 C/Species: Agrobacterium tumefaciens  
 C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
 A/Accession: AF2632  
 R/Wood, D.W.; Seculab, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woc, I  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.  
 A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A/Reference number: AB2577; M0ID:21608550; PMID:11743193  
 A/Accession: AF2632  
 A/Status: Preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-595 <K0M>  
 A/Cross-References: UNIPROT:Q8U143; UNIPARC:UP100000017PD; GB:AE008668; P1DN:AL41476.1;  
 A/Experimental source: strain C58 (Dupont)  
 A/Genetic8:  
 A/Gene: Atu0457  
 A/Map position: circular chromosome

Alignment Scores:  
 Pred. No.: 6.25 Length: 595  
 Score: 96.00 Matches: 74  
 Percent Similarity: 33.9% Conservative: 33  
 Best Local Similarity: 23.4% Mismatches: 99  
 Query Match: 6.2% Indels: 110  
 DB: 2 Gaps: 18

US-10-537-002-7 (1-786) x AF2632 (1-595)

QY 7 GTGACTGCTGTCAAGGCTTGGCTTCA-----CTGATT 48  
 DB 250 ValAlaSerAlaGlnGlyPheSerGlySerValSerAenLysIleLeuArgThrIleAap 269  
 QY 49 GGGATGGCGGCGATCATTTCTGCCACC--TGCATGACCACTGAGACCCCA----- 99  
 DB 270 GlyValAlaArgThrIleAlaThrGlyGlyMetLysGluPheMetAapLeuGly 289  
 QY 100 GACTGTTCACAACACCCCGTACA-----GCTGTTTCAAC 135  
 DB 230 AapLeuLeuLeuGlyProValAapProAenAlaAapArgIleGlyAlaIlePheMet 309  
 QY 136 TACCAAGGCGCTGTGCGCTCTGTGTCGAGAG----- 168  
 DB 310 ArgAlaLysGluTrpGlyAlaSerIleArgGluLeuAenAapAlaIleLysAapAenPro 329  
 QY 169 ---AGCTCTGCTTCAACCGAGTCCGCGGCTAC--TTCACCCCTG-----CTG 210  
 DB 330 IleAlaLysPhePheAlaGluMetSerGlyTrpGlyPheGlnLeuPheAlaTrpGlyMet 349  
 QY 211 GGGCTGCGACGATCTGACGAGGAGTGGCA-----GCCCTGATGATCGA--- 255  
 DB 350 GlyIleSerMetLeuAlaGlyThrIleArgLysLeuAlaAlaLeuAenAapValLeuSer 369  
 QY 256 -----GGCATGCTCCGTGGTGCATTTGGCTCTGTGATTCATCTTGGCTCCCTG----- 303  
 DB 370 GlyAlaSerThrLeuLeuGlyAlaLeuLysThrValGlySerIleAlaIleAlaIleValGly 389  
 QY 304 -----AAATGATCCGATGCGACGATGAGGACCTTGGCAAA 342  
 DB 390 GlyGlyAlaValAlaGlyGlyAlaAlaAlaGlyAlaGlyThrGlySerAlaGlyAlaAla 409

QY 343 GCCAACATGACATGACTCGCGGATCATG-----TTCATTGCTCGACTCT 390  
 DB 410 AlaGlyAlaAlaAlaThrAlaGlyLeuLeuGlyArgTrpGlySerIleIleThrIleLeu 429  
 QY 391 TGTGCATT-----GCTGAGTGTCTGTGTGGCCACATGCTGTGACT 435  
 DB 430 AlaArgIleGlyThrTrpGlyThrAlaGlyAlaGlyLeu----- 442  
 QY 436 AACCTTGGATGTCACAGCTTAACATGTACACCGGCATGGGTGGATGTGACACTGT 495  
 DB 443 -----TTPgluMetGlyLysAlaThrTrpGly----- 452  
 QY 496 CAGACCAAGTACACATTTGGTGGCTGTCTGTGCGGCTGGTGGTGGAGCTTCA 555  
 DB 453 AapThrPheTrpLysGlnGlyLysAla-----TTPLeuProGly----- 465  
 QY 556 CTAAATGGGGGTGTGATGATGTGATCGCTGCCGCGGCGCTGGACACGAAGAACAAC 615  
 DB 466 -----ProGluAapLeuLeu 470  
 QY 616 TCAAAAGCGCTTCTTATACGCTCAGGCCACAGTGTGCTTACAAAGCTTGGAGCTTC 675  
 DB 471 HisAlaValAlaGlySerTrpPheAlaSerSerProSerThrAlaLeuProAlaSerAla 490  
 QY 676 AAGCC-----AGCAGCTGGCTTGGTCCAC 702  
 DB 491 SerAlaGlnLeuSerLeuAbaArgSerAlaArgAlaAlaGlyPheGlyGlySer 510

QY 703 ACCAAAAAC-----AAGAAGATATACGATGAGGTGCCGACCA 741  
 DB 511 ThrAapLeuLeuProGlyLysThrMetAapAapLeuGlyIleArgThr 526

## RESULT 46

QJ01584  
 genome polypotein - hepatitis C virus (strain U.K.) (fragment)  
 N/Contains: core protein C; envelope protein E1; envelope protein E2; nonstructural prote  
 C/Species: hepatitis C virus  
 C/Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004  
 A/Accession: QJ01584  
 R/Kumar, U.; Cheng, D.; Thomas, H.; Montardino, J.  
 J. Gen. Virol. 73, 1521-1525, 1992  
 A/Title: Cloning and sequencing of the structural region and expression of putative core  
 A/Reference number: QJ01584; MUID:92300349; PMID:1318944  
 A/Accession: QJ01584  
 A/Molecule type: genomic RNA  
 A/Residues: 1-640 <K0M>  
 A/Cross-References: UNIPROT:Q68966; UNIPARC:UP100000EF23; GB:X84079; M1D:9643119; P1DN:C  
 A/Superfamily: hepatitis C virus genome polypotein  
 C/Keywords: core protein; envelope protein; glycoprotein; nonstructural protein; polyprot  
 F.1-191/Product: core protein C #status predicted <CPC>  
 F.192-389/Product: envelope protein E1 #status predicted <BE1>  
 F.390-640/Product: envelope protein E2 and nonstructural protein NS1 #status predicted <I  
 F.196,209,234,305,417,430,448,476,540,556,576,623/Binding site: carbohydrate (Aen) (coval

## Alignment Scores:

Pred. No.: 6.21 Length: 640  
 Score: 96.00 Matches: 46  
 Percent Similarity: 38.9% Conservative: 31  
 Best Local Similarity: 23.2% Mismatches: 79  
 Query Match: 6.2% Indels: 42  
 DB: 2 Gaps: 9

US-10-537-002-7 (1-786) x QJ01584 (1-640)

QY 175 GAGTTCACCGAGTGGCGGCTTACTTCACTCTGCGGGCTGCCA----- 219  
 DB 129 GlyPheAlaAapLeuMetGlyTrpIleProLeuValAlaGlyAlaAla 148  
 QY 220 ---GCCATGCTGACGAGCGAGCGCGCTGATGATC--GTAGGATGCTCCTGGTGGCC 273  
 DB 149 ArgAlaLeuAlaHisGlyValArgValLeuGluAapGlyValAenTrpAlaThrGlyAen 168



```

QY 310 ATCCGATTTGGAGCATGAGAGACTCTGCGAAGCCAAATGACACTGACCTCCGGGATC 369
   |||||
Db 163 AsnlyrIleGlySerMet-----AsnPhenylrIleuMetAsnPhen 176
   |||||
QY 370 ATGTCATTTGCTCAGGTCCT-----TGTCGAAATTCGTGAGTGCTGTGTTGGCAAC 423
   |||||
Db 177 MetPheAsnTrpAspLeuLeuTrpPheCysIleuLeuCysAlaPheLeuValIlyMetPro 196
   |||||
QY 424 ATGCTGGAGCTACTTCTTGATGTCACAGCTAACATGATACACGGGATGGGTGG--- 480
   |||||
Db 197 MetPheLeuValHisLeuTrpLeuProIlyValHisValGluAlaProValSerIlySer 216
   |||||
QY 480 ----- 480
Db 217 MetIleLeuAlaGlyIleMetLeuIlyLeuGlyIlyTrpGlyMetLeuArgValIleSer 236
QY 481 ATGTCGACAGACTGTTTCAGACAGTACACATTTGGTGCGGCTCTGTTCTGGGCTGGGTC 540
   |||||
Db 237 PheLeuGlnLeuMetAsnLeuIlySerPhe-----ValTrpIle 250
QY 541 GCTGGAGGCTTCACACTAATTTGGGGTGTGATGATGTGCATCGCTCCGGGGCTGGCA 600
   |||||
Db 251 Ser-----IleSerLeuValGlyIlyValLeuValSerLeuValCysIleu----- 265
QY 601 CCAGAAAGAACCACTACAAAGCC--GTTTCTTATCATGCTCAGGCGCAC 648
Db 266 ---ArgGlnThrAspLeuIlyValLeuIleAlaTrpSerSerValAlaHis 281

RESULT 49
125797
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - fruit fly (Drosophila yakuba) m
N.Alternate names: NADH-ubiquinone oxidoreductase chain 4
C.Species: mitochondrion Drosophila yakuba
C.Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 09-Jul-2004
C.Accession: J25797
R.Clary, D.O.; Wolstenholme, D.R.
J.Mol. Evol. 22, 252-271, 1985
A.Title: The mitochondrial DNA molecule of Drosophila yakuba: nucleotide sequence, gene
A.Reference number: A92962; MUID:86089137; PMID:3001325
A.Accession: J25797
A.Molecule type: DNA
A.Residues: 1-446 <CL>
A.Cross-references: UNIPROT:P07707; UNIPARC:UPI000013075D; GB:X03240; GB:J01400; GB:J014
C.Genetics:
A.Gene: FlyBase:Dyak/mc:ND4
A.Cross-references: FlyBase:FBgn0013185
A.Genome: mitochondrion
A.Genetic code: SGC4
C.Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C.Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Alignment Scores:
Pred. No.: 7.03 Length: 446
Score: 95.50 Matches: 38
Percent Similarity: 41.2% Conservative: 35
Best Local Similarity: 21.5% Mismatches: 57
Query Match: 6.1% Indels: 47
DB: 2 Gaps: 9

US-10-537-002-7 (1-786) x J25797 (1-446)
QY 199 TTCACCTCTGCTGGG-----CTGCCAGCATCTGCAGAGCATGCGAGCCCTGAG 249
   |||||
Db 125 PheLeuIleLeuGlyTrpGlyTrpGlnTrpGlnArgLeuGlnHisIleValIlyrIleuLeu 144
   |||||
QY 250 ATCGTAGGCGATCTCTGGGTGCATTTGGCTCTGATTCATCTTTTCCCTGAATGC 309
   |||||
Db 145 PheTrpTrpIleuLeuVal---SerLeuProMetLeuIleGlyIlyPheTrpVal----- 161
QY 310 ATCCGATTTGGAGCATGAGAGACTCTGCGAAGCCAAATGACACTGACCTCCGGGATC 369
Db 162 -----MetAsnlyrThrGlySerMetAsnPhenylrIleuMetAsnPhen 176

```

```

QY 370 ATGTCATTTGCTCAGGTCCT-----TGTCGAAATTCGTGAGTGCTGTGTTGGCAAC 423
   |||||
Db 177 MetPheAsnTrpAspLeuLeuTrpPheCysIleuLeuCysAlaPheLeuValIlyMetPro 196
   |||||
QY 424 ATGCTGGAGCTACTTCTTGATGTCACAGCTAACATGATACACGGGATGGGTGG--- 480
   |||||
Db 197 MetPheLeuValHisLeuTrpLeuProIlyValHisValGluAlaProValSerIlySer 216
   |||||
QY 480 ----- 480
Db 217 MetIleLeuAlaGlyIleMetLeuIlyLeuGlyIlyTrpGlyMetLeuArgValIleAsn 236
QY 481 ATGTCGACAGACTGTTTCAGACAGTACACATTTGGTGCGGCTCTGTTCTGGGCTGGGTC 540
   |||||
Db 237 PheLeuGlnLeuMetAsnLeuIlySerPhe-----ValTrpIle 250
QY 541 GCTGGAGGCTTCACACTAATTTGGGGTGTGATGATGTGCATCGCTCCGGGGCTGGCA 600
   |||||
Db 251 Ser-----IleSerLeuValGlyIlyValLeuValSerLeuValCysIleu----- 265
QY 601 CCAGAAAGAACCACTACAAAGCC--GTTTCTTATCATGCTCAGGCGCAC 648
Db 266 ---ArgGlnThrAspLeuIlyValLeuIleAlaTrpSerSerValAlaHis 281

RESULT 50
A25704
Synapsin I - rat
C.Species: Rattus norvegicus (Norway rat)
C.Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 09-Jul-2004
C.Accession: A25704
R.McCaferly, C.A.; DeGennaro, L.J.
EMBO J. 5, 3167-3173, 1986
A.Title: Determination and analysis of the primary structure of the nerve terminal speci
A.Reference number: A25704; MUID:87133474; PMID:3028773
A.Residues: 1-691 <MC>
A.Cross-references: UNIPROT:P09951; UNIPARC:UPI0000170812; GB:X04655; NID:g57181; PIDN:C
A.Experimental source: brain
C.Keywords: actin binding; alternative splicing; phosphoprotein

Alignment Scores:
Pred. No.: 6.78 Length: 691
Score: 95.50 Matches: 71
Percent Similarity: 30.7% Conservative: 36
Best Local Similarity: 20.3% Mismatches: 117
Query Match: 6.1% Indels: 125
DB: 2 Gaps: 12

US-10-537-002-7 (1-786) x A25704 (1-691)
QY 729 TTCATCGTATATCTTCTT-----GTTTGGTGCT 700
   |||||
Db 138 SerIleLeuTrpLeuMetProMetValAspSerProTrpTrpIlyPhePheGlyMet 157
   |||||
QY 699 GGACCCAAAGCCAGTGTGAGCTTCAGAGCTTCAGAGCTTGTGTGAGCAACATGTGGCTGA 640
   |||||
Db 158 GlySerIlyValValArgSerLeuIlyProAspPheValIleAlaGlnHisAlaPhe 177
   |||||
QY 639 GGCATGATTAAGAAAGCG-----TTTGTAGTTGGTTTCTTCTGTCGCAAGCC 592
   |||||
Db 178 SerMetAlaArgAsnGlyAspTrpArgSerLeuValIleGlyLeuGlnIlyrAlaIlyle 197
   |||||
QY 591 CCGGAGGCGAGTACATCATTCACACCCCAATTTGTGTGAG-----GCC 547
   |||||
Db 198 ProSerValAsnSerLeuHisSerValIlyrAsnPheCysAspIlySerProTrpValAlaPheAla 217
   |||||
QY 546 TCCAGGACCCAGCCAGCAAGAGCCGACCAATGTGTGATCTGTCTGAACAGTCTG 487
   |||||
Db 218 GlnMetValArgLeuHisIlyIlyValIlyTrpGlnGlnPheProLeuIleAspGlnThr 237
QY 486 CACCATCCACCACCATGCGGCTGATCATGTTAGCTGTGACATCCAGAAATTAGTACAG 427

```

D5	238	PhenylProteinHisIlyeGIuMetLeuSerSerThrThyTrpProValValIlyMet	257
OY	426	CATGTGGCAACACAGACACTCTCCAGCAATTGGACAAAGACCTGAGCAATGACAATGAT	367
D5	258	-----GlyHisAlaHisSerGIuMetGIlyValValIlyValIlySerAlaHisGlnHisAsp	274
OY	367	-----	367
D5	275	PhenAlaPheAlaSerValValAlaLeuThrIlyThrIlyAlaThrAlaGluProPhe	294
OY	366	-----CCGAGAGTGAAGTGCATGTGGCTTTGGC	337
D5	295	IleAspAlaIlyTrpAspValAlaArgValGlnIlyIleGIyGlnIlyTrpValAlaIlyMet	314
OY	336	AGAGCTCCACAGCTGGCAAT	316
D5	315	ArgThrSerValSerGIyAsnTrpIlyThrAsnThrGIySerAlaMetLeuGlnIle	334
OY	316	-----	316
D5	335	AlaMetSerAspArgIlyIlyLeuTrpValAlaAspThrCysSerGIuIlePheGlyGIyLeu	354
OY	315	-----GGGAGTCAATTTACGGCAAAAGATGA-----TACCAAGAG	280
D5	355	AspIleCysAlaValaGlnAlaLeuHisGIyIlyAspGIyAlaArgAspHisIleIleGluVal	374
OY	279	GCCATGGCACCCAGAGACATGCCATACATCATCAGAGGCTCGACATGCCCTCGACATGGC	220
D5	375	ValGIySerSerMetProLeuIleGIyAspHisGlnGIyAspGIyAspIlySerGlnLeuIle	394
OY	219	TGGCAG-----CCCCAGCAGGGGTGAAGTAGCCCG	190
D5	395	ValGIuLeuValValAsnIlyMetThrGlnAlaLeuProArgGlnArgAspAlaSerPro	414
OY	189	GCACTCGGTGAAGCCAGACCTCTCGACACACAGAGCGCCACAGCCCTGTAGTTGAA	130
D5	415	GIyArgGly--Ser-----HisSerGlnThrProSerProGIyAla----	427
OY	129	AACAGCTGTACGGGGTTGTGTACAAGTCTTGAGTCCACT-----G	85
D5	428	-----LeuProLeuGIyArgGlnThrSerGlnIleProAlaG	440
OY	84	GTCCATGCAAGTGGGAGCAATGATGCCCGCAATCCCAATCAGTGAACCAACGAAACCCCA	25
D5	440	IlyProProAlaGlnGlnGlnIlyProProProGIuGlnGIyIlyProGIuInPro--GIyProG	459
OY	24	GGCCTGACAGGACGTACGGCCA	2
D5	459	IlyProGlnArgGlnGIyProPro	466

Search completed: September 26, 2006, 00:22:03  
Job time : 80.5 secs

**THIS PAGE IS BLANK**



This Page is inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images  
problems checked, please do not report the  
problems to the IFW Image Problem Mailbox**

**THIS PAGE IS BLANK**